

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2006, 05:46:47 ; Search time 39.2079 Seconds
(without alignments)
69.968 Million cell updates/sec

Title: US-10-541-343-2

Perfect score: 31

Sequence: 1 FSRYPAR 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 500 summaries

Database :

1: A_Geneseq.8:*
2: geneseqp1980s:*
3: geneseqp1990s:*
4: geneseqp2000s:*
5: geneseqp2001s:*
6: geneseqp2002s:*
7: geneseqp2003as:*
8: geneseqp2004s:*
9: geneseqp2005s:*
10: geneseqp2006s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	100.0	6	ADQ90196	Adq90196 FSRYPAR 6
2	31	100.0	10	ADQ90198	Adq90198 MNTF 10 m
3	31	100.0	11	ADQ90199	Adq90199 MNTF 11 m
4	31	100.0	13	ADQ90200	Adq90200 MNTF 13 m
5	31	100.0	21	ADQ90201	Adq90201 MNTF 21 m
6	31	100.0	9	AEAI7675	Aeai7675 Motoneuro
7	31	100.0	33	AAW59046	AAW59046 Human MNT
8	31	100.0	33	AAO29914	AAO29914 Human mot
9	31	100.0	33	ADQ90195	Adq90195 MNTF1-P6
10	31	100.0	81	ABMS4275	Abms4275 Propionib
11	31	100.0	81	AAU57258	Aau57258 Propionib
12	31	100.0	112	ABMS3777	Abms3777 Propionib
13	31	100.0	112	ABMS4633	Abms4633 Propionib
14	31	100.0	357	AAV35559	Aav35559 Chlamydia
15	31	100.0	357	ABR90579	Abbr90579 Chlamydia
16	31	100.0	357	ABU27032	Abu27032 Protein e
17	31	100.0	357	ABU27032	Abu27032 Protein e
18	31	100.0	357	ABU27032	Abu27032 Protein e
19	31	100.0	463	ADSA4828	Adsa4828 Bacterial
20	31	100.0	466	AAE23630	Aae23630 Escherich
21	31	100.0	474	AAU34677	Aau34677 E. coli c
22	31	100.0	474	ADSA5088	Adsa5088 Bacterial
23	31	100.0	509	ABO74595	Abot74595 Pseudomon

24	31	100.0	1225	4	ABG24444	Abg24444 Novel hum
25	31	100.0	1258	4	ABG24819	Abg24819 Novel hum
26	28	90.3	58	4	AAU60688	Aau60688 Propionib
27	28	90.3	58	6	ABMS7207	Abms7207 Propionib
28	28	90.3	85	2	AAV12203	Aav12203 Human 5'
29	28	90.3	90	6	ABR40679	Abrr40679 Oryza sat
30	28	90.3	132	2	ABR40677	Abrr40677 Oryza sat
31	28	90.3	139	2	AAW67901	Aaw67901 Human sec
32	28	90.3	155	6	ADB06538	Adb06538 Alloiococ
33	28	90.3	172	6	ADB06540	Adb06540 Alloiococ
34	28	90.3	174	2	AAW67842	Aaw67842 Human sec
35	28	90.3	174	6	ADY14896	Ady14896 PRO polyp
36	28	90.3	174	10	AEF69875	Aef69875 Microsate
37	28	90.3	243	6	AEAI7517	Aeai7517 Thale cre
38	28	90.3	244	6	ABR40670	Abrr40670 Zea mays
39	28	90.3	256	8	ADX91903	Adx91903 Plant ful
40	28	90.3	301	4	ABG07163	Abg07163 Novel hum
41	28	90.3	301	9	AEAI7523	Aeai7523 Thale cre
42	28	90.3	301	9	AEAI7522	Aeai7522 Thale cre
43	28	90.3	499	4	ABG27724	Abg27724 Novel hum
44	28	90.3	520	4	ABG29681	Abg29681 Novel hum
45	28	90.3	661	8	ADN18185	Adn18185 Bacterial
46	28	90.3	847	4	ABG04036	Abg04036 Novel hum
47	28	90.3	1206	6	ABP76771	Abp76771 N. gonorr
48	28	90.3	1206	6	ABP76681	Abp76681 N. gonorr
49	28	90.3	1718	7	ABG69413	Abg69413 Pseudomon
50	28	90.3	2151	5	ABR78221	Abrr78221 Alpha1C s
51	27	87.1	51	8	ADS07116	Ads07116 Stephyloc
52	27	87.1	59	4	AAU14729	Aau14729 Novel bon
53	27	87.1	182	7	ABM89390	Abm89390 Rice abio
54	27	87.1	198	7	ABO83305	Abos83305 Pseudomon
55	27	87.1	208	3	AAG22027	Aag22027 Arabidops
56	27	87.1	209	3	AAG46509	Aag46509 Arabidops
57	27	87.1	209	3	AAG53926	Aag53926 Arabidops
58	27	87.1	209	3	AAG24637	Aag24637 Arabidops
59	27	87.1	209	8	ADT56063	Adt56063 Plant pol
60	27	87.1	217	3	AAG24636	Aag24636 Arabidops
61	27	87.1	217	3	AAG46508	Aag46508 Arabidops
62	27	87.1	218	3	AAG22026	Aag22026 Arabidops
63	27	87.1	220	3	ABAB3764	Abab3764 Human can
64	27	87.1	272	8	ADX90400	Adx90400 Plant ful
65	27	87.1	284	3	AAG05948	Aag05948 Arabidops
66	27	87.1	290	3	AAG05372	Aag05372 Arabidops
67	27	87.1	313	3	AAG05947	Aag05947 Arabidops
68	27	87.1	341	2	AAW54385	Aaw54385 Actinomad
69	27	87.1	354	4	AAU36201	Aau36201 Pseudomon
70	27	87.1	354	6	ABU38270	Abu38270 Protein e
71	27	87.1	354	6	ABU41318	Abu41318 Protein e
72	27	87.1	364	7	ABO71135	Abot71135 Pseudomon
73	27	87.1	387	7	ABO67510	Abot67510 Klebsiell
74	27	87.1	409	8	ADSA1920	Adsa1920 Bacterial
75	27	87.1	410	4	ABAB96590	Abab96590 Putative
76	27	87.1	410	8	ADN18631	Adn18631 Bacterial
77	27	87.1	410	8	ADSA3194	Adsa3194 Bacterial
78	27	87.1	414	7	ADBE80738	Adbe80738 Microsate
79	27	87.1	428	7	ADBE80739	Adbe80739 Microsate
80	27	87.1	445	3	AAAB16689	Aaab16689 Bacteriop
81	27	87.1	453	8	ADS25032	Ads25032 Bacterial
82	27	87.1	456	5	AAE23631	Aae23631 Escherich
83	27	87.1	456	6	ABU47476	Abu47476 Protein e
84	27	87.1	456	6	ABU46912	Abu46912 Protein e
85	27	87.1	457	6	ABU32173	Abu32173 Protein e
86	27	87.1	457	6	ABU15160	Abu15160 Protein e
87	27	87.1	457	8	ADN18076	Adn18076 Bacterial
88	27	87.1	460	2	AAV35220	Aav35220 Chlamydia
89	27	87.1	460	7	ABO66988	Abos66988 Klebsiell
90	27	87.1	463	4	AAU38325	Aau38325 Salmoneil
91	27	87.1	463	6	ABU50389	Abu50389 Protein e
92	27	87.1	463	8	ADN17661	Adn17661 Bacterial
93	27	87.1	469	7	ABO61707	Abos61707 Klebsiell
94	27	87.1	478	8	ADSA27797	Adsa27797 Bacterial
95	27	87.1	483	8	ADSA25210	Adsa25210 Bacterial
96	27	87.1	484	7	ADD22903	Add22903 T. reesei

97	27	87.1	484	9	AED25788	Aed25788	Trichoder	170	26	83.9	123	4	AAU02628	Au02628	Anti-adip
98	27	87.1	488	5	AAU79550	Aau79550	A. naesli	171	26	83.9	123	9	AED8738	B6 heavy	
99	27	87.1	491	6	ABU41654	Abu41654	Protein e	172	26	83.9	124	9	AD245429	Murine fa	
100	27	87.1	505	8	ADS25873	AdS25873	Bacterial	173	26	83.9	125	2	AAW28065	Staphyloc	
101	27	87.1	505	8	ADS25472	AdS25472	Bacterial	174	26	83.9	125	2	AAV39818	Llama ant	
102	27	87.1	538	5	ABP39788	Abp39788	Staphyloc	175	26	83.9	125	7	ABO33840	Human ant	
103	27	87.1	538	8	ADSO5529	AdSO5529	Staphyloc	176	26	83.9	125	9	AD235786	Anti-glic	
104	27	87.1	545	8	ADX75725	AdX75725	Plant ful	177	26	83.9	126	9	AED63549	Heavy cha	
105	27	87.1	575	8	ADR10067	AdR10067	Human pro	178	26	83.9	127	8	ADO39423	Human AB7	
106	27	87.1	579	4	AAE12776	Aae12776	Chlamydia	179	26	83.9	128	4	AAAG71332	Human sec	
107	27	87.1	592	8	ADJ10121	AdJ10121	Rat vitam	180	26	83.9	128	4	AAAG71342	Human sec	
108	27	87.1	607	8	ADS22813	AdS22813	Bacterial	181	26	83.9	128	4	AAAG71346	Human sec	
109	27	87.1	607	8	ADS23171	AdS23171	Bacterial	182	26	83.9	128	4	AAAG71348	Human sec	
110	27	87.1	607	8	ADS23197	AdS23197	Bacterial	183	26	83.9	128	4	AAAG71336	Human sec	
111	27	87.1	608	8	ADS22510	AdS22510	Bacterial	184	26	83.9	128	4	AAAG71334	Human sec	
112	27	87.1	630	4	AAW67683	Aaw67683	Corynebac	185	26	83.9	128	4	AAAG71338	Human sec	
113	27	87.1	630	4	AAW92305	Aaw92305	C. glutam	186	26	83.9	128	4	AAAG71352	Human sec	
114	27	87.1	635	7	ADL65941	AdL65941	C. glutam	187	26	83.9	128	4	AAAG71344	Human sec	
115	27	87.1	635	7	ADL65563	AdL65563	C. glutam	188	26	83.9	128	4	AAAG71344	Human sec	
116	27	87.1	647	6	ABM67463	Abm67463	Phototrab	189	26	83.9	129	9	ADX02195	SARS coro	
117	27	87.1	650	7	ADE80737	AdE80737	Microsate	190	26	83.9	130	2	AAW71045	Peptide s	
118	27	87.1	650	8	ADJ10123	AdJ10123	Pig vitam	191	26	83.9	130	10	AEG01402	Kalikkrei	
119	27	87.1	650	8	ADJ10117	AdJ10117	Human vit	192	26	83.9	131	10	AEG01403	Kalikkrei	
120	27	87.1	661	9	AEC35957	Aec35957	Varicella	193	26	83.9	137	6	ABU97068	Recombina	
121	27	87.1	672	5	AAU75887	Aau75887	Human adh	194	26	83.9	140	6	ABG74241	Mouse ant	
122	27	87.1	675	5	ABO80415	AbO80415	Pseudomon	195	26	83.9	142	10	AEG01479	Kalikkrei	
123	27	87.1	912	7	ABM86634	Abm86634	Rice abio	196	26	83.9	142	10	AEG01504	Kalikkrei	
124	27	87.1	1243	6	ABG99964	Abg99964	Human nov	197	26	83.9	143	4	AAAG71354	Human sec	
125	27	87.1	1377	6	ABP56518	Abp56518	Human BAA	198	26	83.9	145	6	ABP78858	N. gonorr	
126	27	87.1	1421	6	ABO52987	AbO52987	Human put	199	26	83.9	146	5	ABP05286	Human ORF	
127	27	87.1	1556	6	ABO53101	AbO53101	Human put	200	26	83.9	164	9	AAAG76005	Human col	
128	27	87.1	1556	6	ADY25712	ADY25712	MPREN AQR	201	26	83.9	174	8	ADN19577	Bacterfal	
129	27	87.1	2190	7	ADB80326	AdB80326	Human MDR	202	26	83.9	189	6	ABP78997	N. gonorr	
130	26	83.9	10	5	AAU82535	Aau82535	Llama CDR	203	26	83.9	198	4	ABG02424	Novel hum	
131	26	83.9	10	5	AAU82532	Aau82532	Llama CDR	204	26	83.9	198	6	ABP99711	Human sec	
132	26	83.9	28	1	AAE80205	Aae80205	Sequence	205	26	83.9	198	6	ABR01192	Human gen	
133	26	83.9	49	6	ABP77874	Abp77874	N. gonorr	206	26	83.9	198	7	ADC20475	Human sec	
134	26	83.9	50	8	ADH117396	AdH117396	Human NOV	207	26	83.9	199	3	AAAB29815	Human sec	
135	26	83.9	104	4	AAU86940	Aau86940	Human DNA	208	26	83.9	203	2	AAW28268	Amno aci	
136	26	83.9	104	4	AAU87614	Aau87614	Novel cen	209	26	83.9	219	3	AAAG40819	Zea mayr	
137	26	83.9	104	4	ADL54929	AdL54929	Novel hum	210	26	83.9	229	3	AAV53611	The nitri	
138	26	83.9	107	3	ABW41543	AbW41543	Human ORF	211	26	83.9	232	3	AAAB29870	Human sec	
139	26	83.9	108	7	ADF06021	AdF06021	Bacterial	212	26	83.9	233	7	AAO66201	Klebsiell	
140	26	83.9	111	7	ADF05922	AdF05922	Bacterial	213	26	83.9	247	1	AAAB08155	Frs and C	
141	26	83.9	117	4	AAU02590	Aau02590	Anti-adip	214	26	83.9	247	1	AAAP80156	Biosynthe	
142	26	83.9	118	2	AAV05237	Aav05237	Testoster	215	26	83.9	247	5	ABP45317	Human Bly	
143	26	83.9	118	2	AAV05238	Aav05238	Testoster	216	26	83.9	247	5	ABP44867	Human Bly	
144	26	83.9	118	2	AAV05235	Aav05235	Testoster	217	26	83.9	247	5	ABP445088	Human Bly	
145	26	83.9	118	2	AAV05236	Aav05236	Testoster	218	26	83.9	247	5	ABP45079	Human Bly	
146	26	83.9	118	4	AAU02530	Aau02530	Anti-adip	219	26	83.9	247	7	ADG95694	Single ch	
147	26	83.9	119	2	AAW92998	Aaw92998	Humanised	220	26	83.9	247	7	ADG95906	Single ch	
148	26	83.9	119	2	AAW92996	Aaw92996	Monoclona	221	26	83.9	247	7	ADG96144	Single ch	
149	26	83.9	119	2	AAW29996	Aaw29996	Humanised	222	26	83.9	247	7	ADG95915	Single ch	
150	26	83.9	119	2	AAW29994	Aaw29994	Heavy cha	223	26	83.9	247	8	ADO04878	Anti-mous	
151	26	83.9	119	9	AEC76981	Aec76981	Protein e	224	26	83.9	247	8	ADP09937	Anti-mous	
152	26	83.9	119	9	AEC76980	Aec76980	Protein e	225	26	83.9	247	9	AED77747	Human B L	
153	26	83.9	119	10	AEG08773	Aeg08773	Tie recep	226	26	83.9	247	9	AED77968	Human B L	
154	26	83.9	120	8	ADP09916	Adp09916	Anti-huma	227	26	83.9	247	9	AED77959	Human B L	
155	26	83.9	120	8	ADP09917	Adp09917	Anti-huma	228	26	83.9	247	9	AED78197	Human B L	
156	26	83.9	120	8	AD025752	Ad025752	Anti-IFN	229	26	83.9	249	4	AAU35287	Enterococ	
157	26	83.9	120	8	AD025751	Ad025751	Anti-IFN	230	26	83.9	249	4	AED30836	Llama ant	
158	26	83.9	120	8	AD025875	Ad025875	Camelidae	231	26	83.9	251	3	AAAB43605	Human can	
159	26	83.9	120	8	AD025876	Ad025876	Camelidae	232	26	83.9	252	7	ABM85609	Human pro	
160	26	83.9	120	10	AEG08921	Aeg08921	Tie recep	233	26	83.9	252	9	AD270532	Human pro	
161	26	83.9	121	4	AAU02598	Aau02598	Anti-adip	234	26	83.9	253	5	ABP45188	Human Bly	
162	26	83.9	121	5	ABP30699	Abp30699	Streptoco	235	26	83.9	253	7	ADG96015	Single ch	
163	26	83.9	121	5	ABG61524	Abg61524	Iron upla	236	26	83.9	253	9	AED78068	Human B L	
164	26	83.9	121	6	ABU01471	AbU01471	S. pneumo	237	26	83.9	255	5	ABP45001	Human Bly	
165	26	83.9	121	8	ADV88642	Adv88642	Streptoco	238	26	83.9	255	7	ADG95828	Single ch	
166	26	83.9	121	8	ADV82041	Adv82041	Streptoco	239	26	83.9	255	9	AED77881	Human B L	
167	26	83.9	121	8	ADV79895	Adv79895	Streptoco	240	26	83.9	256	9	ADX02030	SARS coro	
168	26	83.9	122	4	AAG71350	Aag71350	Human sec	241	26	83.9	264	2	AAV34848	Chlamydia	
169	26	83.9	122	9	AED30809	Aed30809	Llama ant	242	26	83.9	264	6	ABU97152	Recombina	

243	26	83.9	267	6	ABU29331	Abu29331 Protein e	316	26	83.9	473	4	AAG64471	Aag64471 Human typ
244	26	83.9	271	3	AAG34202	Aag34202 Zee may	317	26	83.9	473	4	AAG64469	Aag64469 Human typ
245	26	83.9	271	3	AAG40818	Aag40818 DNA may	318	26	83.9	473	4	AAG64473	Aag64473 Human typ
246	26	83.9	271	7	ABM74308	Abm74308 Cione	319	26	83.9	475	7	ABOC3637	Aboc3637 Klebsiell
247	26	83.9	272	9	ADZ58597	Adz58597 Novel sin	320	26	83.9	476	9	AED19758	Aed19758 Chimeric
248	26	83.9	276	7	ABO76413	AbO76413 Pseudomon	321	26	83.9	495	4	ABG09718	Abg09718 Novel hum
249	26	83.9	279	8	ADX88115	Adx88115 Plant ful	322	26	83.9	495	4	AAH89668	Aah89668 DRCl prot
250	26	83.9	283	7	ABO71739	AbO71739 Pseudomon	323	26	83.9	497	4	AAU37751	Aau37751 Streptoco
251	26	83.9	285	7	ABM5608	Abm5608 Mouse pro	324	26	83.9	497	6	ABU01284	Abu01284 S. pneumo
252	26	83.9	299	7	ADV21133	Adv21133 Human pro	325	26	83.9	497	6	ABU45988	Abu45988 Protein e
253	26	83.9	303	8	ADP98971	Adp98971 C. albica	326	26	83.9	497	8	ADK48573	Adk48573 Streptoco
254	26	83.9	307	8	ADO07053	Ado07053 Human pro	327	26	83.9	497	8	ADM92139	Adm92139 S. pneumon
255	26	83.9	311	3	AAG34200	Aag34200 Zee may	328	26	83.9	504	8	ADR95097	Adr95097 Novel S.
256	26	83.9	315	8	ADN36928	Adn36928 X. albii	329	26	83.9	504	7	AEA58967	Aea58967 Streptoco
257	26	83.9	319	8	ABN36928	Abn36928 X. albii	330	26	83.9	506	7	ABOC1660	Aboc1660 Klebsiell
258	26	83.9	319	4	ABAB36481	Abab36481 Fusion ge	331	26	83.9	509	6	ABJ18916	Abj18916 Pathogen
259	26	83.9	321	4	ABAB67541	Abab67541 Amino aci	332	26	83.9	509	6	ABM72709	Abm72709 Staphyloc
260	26	83.9	327	7	ABO67488	AbO67488 C. neofo	333	26	83.9	520	6	ABJ25847	Abj25847 Aspergill
261	26	83.9	328	3	ABAB36480	Abab36480 Klebsiell	334	26	83.9	520	6	ABM73438	Abm73438 Staphyloc
262	26	83.9	328	3	ABAB36480	Abab36480 Fusion ge	335	26	83.9	528	6	ABJ26447	Abj26447 Aspergill
263	26	83.9	335	4	AAU87326	Aau87326 Novel aci	336	26	83.9	543	6	ABU31756	Abu31756 Protein e
264	26	83.9	335	4	AAU87326	Aau87326 Novel cen	337	26	83.9	544	7	ABO80249	AbO80249 Pseudomon
265	26	83.9	336	4	AD154641	Ad154641 Novel hum	338	26	83.9	547	5	AAE14507	Aae14507 Human car
266	26	83.9	336	5	ABG63444	Abg63444 Human gen	339	26	83.9	547	7	ABO64232	AbO64232 Klebsiell
267	26	83.9	336	5	ABG63444	Abg63444 Human alb	340	26	83.9	547	8	AD179888	Ad179888 Protein o
268	26	83.9	339	6	ABG72018	Abg72018 Albumin f	341	26	83.9	549	7	ADH80159	Adh80159 Mycobacte
269	26	83.9	340	5	ABP39211	Abp39211 Mouse cha	342	26	83.9	557	4	ABG07424	Abg07424 Novel hum
270	26	83.9	340	5	ABP39211	Abp39211 Staphyloc	343	26	83.9	557	7	ADG42192	Adg42192 Human bta
271	26	83.9	341	6	ADAB36377	Ada36377 Staphyloc	344	26	83.9	558	8	ADN20745	Adn20745 Bacteri
272	26	83.9	342	6	ADO28579	Ado28579 Acinetoba	345	26	83.9	571	3	AAV71107	Aav71107 Human Hyd
273	26	83.9	342	8	ADX65968	Adx65968 Plant ful	346	26	83.9	571	3	AAAB18914	Aab18914 A. novel p
274	26	83.9	342	9	AECL1995	Aecl1995 Rat suro	347	26	83.9	571	4	AAU12442	Aau12442 Human PRO
275	26	83.9	342	9	AECL1995	Aecl1995 Rat suro	348	26	83.9	571	4	AAE04101	Aae04101 Human gen
276	26	83.9	342	9	AECL1996	Aecl1996 Rat suro	349	26	83.9	571	5	AAU83696	Aau83696 Human PRO
277	26	83.9	342	10	AEF27704	Aef27704 Rat prost	350	26	83.9	571	5	ABAB84949	Abab84949 Human PRO
278	26	83.9	342	10	AEF27703	Aef27703 Rat prost	351	26	83.9	571	5	ABG64341	Abg64341 Human alb
279	26	83.9	343	5	ABG96402	Abg96402 Human ova	352	26	83.9	571	5	ADY31950	Ady31950 Novel hum
280	26	83.9	343	5	ABU78547	Abu78547 Human pro	353	26	83.9	571	6	ABO17886	AbO17886 Novel hum
281	26	83.9	343	5	ABU78547	Abu78547 Human pro	354	26	83.9	571	6	ABU69108	Abu69108 Human PRO
282	26	83.9	343	6	ABU37067	Abu37067 Humno bre	355	26	83.9	571	6	ABU80843	Abu80843 Human PRO
283	26	83.9	343	6	ABU37067	Abu37067 Humno bre	356	26	83.9	571	6	ABO33809	AbO33809 Novel hum
284	26	83.9	343	7	ADD47565	Add47565 Human Pro	357	26	83.9	571	6	ABU81140	AbU81140 Human PRO
285	26	83.9	343	7	ADD47561	Add47561 Human Pro	358	26	83.9	571	6	ABO19424	AbO19424 Human sec
286	26	83.9	343	7	ADG38955	Adg38955 Human pro	359	26	83.9	571	6	ABU66840	Abu66840 Human PRO
287	26	83.9	343	8	ADN03871	Adn03871 Antipsoxi	360	26	83.9	571	6	ABU59921	Abu59921 Novel sec
288	26	83.9	343	8	ADV25080	Adv25080 Human ser	361	26	83.9	571	6	ABU59085	Abu59085 Human PRO
289	26	83.9	343	9	AECL1215	Aecl1215 Human sur	362	26	83.9	571	6	ABO25111	AbO25111 Human sec
290	26	83.9	343	9	AECL1215	Aecl1215 Human sur	363	26	83.9	571	6	ABU82152	AbU82152 Novel hum
291	26	83.9	343	10	AECL1217	Aecl1217 Human sur	364	26	83.9	571	6	ABU67116	Abu67116 Human sec
292	26	83.9	343	10	AEF27700	Aef27700 Human tar	365	26	83.9	571	6	ABU81549	Abu81549 Human sec
293	26	83.9	348	6	ABU18926	Abu18926 Pathogen	366	26	83.9	571	6	ADA46061	Ada46061 Novel hum
294	26	83.9	348	6	ABM72980	Abm72980 Staphyloc	367	26	83.9	571	6	ADA76548	Ada76548 Novel hum
295	26	83.9	349	8	ADT59250	Adt59250 Plant pol	368	26	83.9	571	6	ADA76492	Ada76492 Human PRO
296	26	83.9	356	8	ADN20668	Adn20668 Bacteri	369	26	83.9	571	6	ABJ72332	Abj72332 Human PRO
297	26	83.9	363	7	ABO84037	AbO84037 Pseudomon	370	26	83.9	571	6	ADA19142	Ada19142 Human PRO
298	26	83.9	365	8	ADOS7564	Ados7564 Actinobac	371	26	83.9	571	6	ADA61765	Ada61765 Homo sapi
299	26	83.9	376	8	ADM91934	Adm91934 M. xanthu	372	26	83.9	571	6	ADB19550	Adb19550 Novel hum
300	26	83.9	386	5	ABE98416	AbE98416 Human NOV	373	26	83.9	571	6	ADB28091	AdB28091 Human PRO
301	26	83.9	393	7	ADF06202	Adf06202 Bacteri	374	26	83.9	571	6	ADA86570	Ada86570 Novel hum
302	26	83.9	393	7	ADF06202	Adf06202 Bacteri	375	26	83.9	571	6	ADB16134	Adb16134 Human PRO
303	26	83.9	431	5	ABBA48085	Abba48085 Human pla	376	26	83.9	571	6	ADA47920	Ada47920 Human PRO
304	26	83.9	435	7	AAAB20163	Aaab20163 Human pro	377	26	83.9	571	6	ADA67715	Ada67715 Human PRO
305	26	83.9	435	7	ADG38848	Adg38848 Human SEC	378	26	83.9	571	6	ADB30722	Adb30722 Human PRO
306	26	83.9	448	6	ABU44929	Abu44929 Protein e	379	26	83.9	571	6	ADA86018	Ada86018 Novel hum
307	26	83.9	450	8	ADS30536	Ads30536 Bacteri	380	26	83.9	571	6	ADA97230	Ada97230 Human PRO
308	26	83.9	451	8	ADH17410	Adh17410 Human NOV	381	26	83.9	571	6	ADA79534	Ada79534 Human PRO
309	26	83.9	451	8	ADH17412	Adh17412 Human NOV	382	26	83.9	571	6	ADA87673	Ada87673 Novel hum
310	26	83.9	451	8	ADH17394	Adh17394 Human NOV	383	26	83.9	571	6	ADB16875	Adb16875 Human PRO
311	26	83.9	463	8	ADM72025	Adm72025 Chimeric	384	26	83.9	571	6	ADA91967	Ada91967 Novel hum
312	26	83.9	463	10	AEF50991	Aef50991 Variable	385	26	83.9	571	6	ADB15030	Adb15030 Human PRO
313	26	83.9	466	1	AAAP40032	Aap40032 Gamma ant	386	26	83.9	571	6	ADB18991	Adb18991 Novel hum
314	26	83.9	467	6	ABU41994	Abu41994 Protein e	387	26	83.9	571	6	ADA94206	Ada94206 Human PRO
315	26	83.9	473	4	AAG64475	Aag64475 Human typ	388	26	83.9	571	6	ADB20102	Adb20102 Novel hum

389	26	83.9	571	6	ADB13414	Adh13414	Human	PRO
390	26	83.9	571	6	ABO43419	AbO43419	Novel	hum
391	26	83.9	571	6	ADA74668	Ada74668	Human	PRO
392	26	83.9	571	6	ADB24901	Adb24901	Human	PRO
393	26	83.9	571	6	ADA82425	Ada82425	Human	PRO
394	26	83.9	571	6	ADA75388	Ada75388	Human	PRO
395	26	83.9	571	6	ADA85466	Ada85466	Novel	hum
396	26	83.9	571	6	ADA84914	Ada84914	Novel	hum
397	26	83.9	571	6	ADB30170	Adb30170	Human	PRO
398	26	83.9	571	6	ADA80698	Ada80698	Human	PRO
399	26	83.9	571	6	ADA75940	Ada75940	Human	PRO
400	26	83.9	571	6	ADA47165	Ada47165	Human	PRO
401	26	83.9	571	6	ADB25461	Adb25461	Human	PRO
402	26	83.9	571	6	ADA93637	Ada93637	Human	PRO
403	26	83.9	571	6	ADB26987	Adb26987	Human	PRO
404	26	83.9	571	6	ADB31274	Adb31274	Human	PRO
405	26	83.9	571	6	ABJ72460	Abj72460	Human	PRO
406	26	83.9	571	6	ADA61202	Ada61202	Homo sapi	
407	26	83.9	571	6	ADB24349	Adb24349	Human	PRO
408	26	83.9	571	6	ADA96678	Ada96678	Human	PRO
409	26	83.9	571	6	ADA81250	Ada81250	Human	PRO
410	26	83.9	571	6	ADA96126	Ada96126	Human	PRO
411	26	83.9	571	6	ADB26435	Adb26435	Human	PRO
412	26	83.9	571	6	ADB21920	Adb21920	Novel	hum
413	26	83.9	571	6	ABO34355	AbO34355	Human	sec
414	26	83.9	571	7	ADA77699	Ada77699	Human	PRO
415	26	83.9	571	7	ADB18439	Adb18439	Human	PRO
416	26	83.9	571	7	ADA87122	Ada87122	Novel	hum
417	26	83.9	571	7	ABO25132	AbO25132	Human	sec
418	26	83.9	571	7	ADA88825	Ada88825	Novel	hum
419	26	83.9	571	7	ADA46613	Ada46613	Novel	hum
420	26	83.9	571	7	ADB28643	Adb28643	Human	PRO
421	26	83.9	571	7	ADB29195	Adb29195	Human	PRO
422	26	83.9	571	7	ADA77147	Ada77147	Human	PRO
423	26	83.9	571	7	ADA88777	Ada88777	Novel	hum
424	26	83.9	571	7	ADA97782	Ada97782	Human	PRO
425	26	83.9	571	7	ADB27539	Adb27539	Human	PRO
426	26	83.9	571	7	ADB22472	Adb22472	Novel	hum
427	26	83.9	571	7	ABJ72162	Abj72162	Human	mem
428	26	83.9	571	7	ADA67163	Ada67163	Human	PRO
429	26	83.9	571	7	ADB23024	Adb23024	Human	PRO
430	26	83.9	571	7	ADB23797	Adb23797	Human	PRO
431	26	83.9	571	7	ADA92519	Ada92519	Novel	hum
432	26	83.9	571	7	ADB15582	Adb15582	Human	PRO
433	26	83.9	571	7	ADB83700	Adb83700	Novel	hum
434	26	83.9	571	7	ADB80806	Adb80806	Novel	hum
435	26	83.9	571	7	ADB73347	Adb73347	Novel	hum
436	26	83.9	571	7	ADB38834	Adb38834	Novel	hum
437	26	83.9	571	7	ADB78429	Adb78429	Novel	hum
438	26	83.9	571	7	ADB38282	Adb38282	Novel	hum
439	26	83.9	571	7	ADB66754	Adb66754	Novel	hum
440	26	83.9	571	7	ADB85077	Adb85077	Human	PRO
441	26	83.9	571	7	ADB89834	Adb89834	Human	PRO
442	26	83.9	571	7	ADB90566	Adb90566	Human	PRO
443	26	83.9	571	7	ADB39667	Adb39667	Novel	hum
444	26	83.9	571	7	ADB78183	Adb78183	Novel	hum
445	26	83.9	571	7	ADB87249	Adb87249	Human	PRO
446	26	83.9	571	7	ADB84831	Adb84831	Human	PRO
447	26	83.9	571	7	ADB47290	Adb47290	Novel	hum
448	26	83.9	571	7	ADB83946	Adb83946	Novel	hum
449	26	83.9	571	7	ADB86897	Adb86897	Human	PRO
450	26	83.9	571	7	ADB73101	Adb73101	Novel	hum
451	26	83.9	571	7	ADB77502	Adb77502	Novel	hum
452	26	83.9	571	7	ADB34659	Adb34659	Human	PRO
453	26	83.9	571	7	ADB35763	Adb35763	Human	PRO
454	26	83.9	571	7	ADB34107	Adb34107	Human	PRO
455	26	83.9	571	7	ADB35211	Adb35211	Human	PRO
456	26	83.9	571	7	ADB36315	Adb36315	Human	PRO
457	26	83.9	571	7	ADB46710	Adb46710	Novel	hum
458	26	83.9	571	7	AAE39102	Aae39102	Human	PRO
459	26	83.9	571	7	AAE39039	Aae39039	Human	PRO
460	26	83.9	571	7	ADC36939	Adc36939	Human	PRO
461	26	83.9	571	7	ADC21929	Adc21929	Human	PRO

462	26	83.9	571	7	ADC50583	Adc50583	Novel	hum
463	26	83.9	571	7	ADC72130	Adc72130	Novel	hum
464	26	83.9	571	7	ADC29779	Adc29779	Novel	hum
465	26	83.9	571	7	ADC60109	Adc60109	Novel	hum
466	26	83.9	571	7	ADC49960	Adc49960	Novel	hum
467	26	83.9	571	7	ADC49159	Adc49159	Novel	hum
468	26	83.9	571	7	ADC49676	Adc49676	Novel	hum
469	26	83.9	571	7	ADC47537	Adc47537	Novel	hum
470	26	83.9	571	7	ADC53116	Adc53116	Novel	hum
471	26	83.9	571	7	ADC57470	Adc57470	Novel	hum
472	26	83.9	571	7	ADC60661	Adc60661	Novel	hum
473	26	83.9	571	7	ADC51136	Adc51136	Novel	hum
474	26	83.9	571	7	ADC65663	Adc65663	Human	PRO
475	26	83.9	571	7	ADC54761	Adc54761	Novel	hum
476	26	83.9	571	7	ADC53722	Adc53722	Novel	hum
477	26	83.9	571	7	ADC59245	Adc59245	Novel	hum
478	26	83.9	571	7	ADC56123	Adc56123	Novel	hum
479	26	83.9	571	7	ADC58693	Adc58693	Novel	hum
480	26	83.9	571	7	ADC47282	Adc47282	Novel	hum
481	26	83.9	571	7	ADC03367	Adc03367	Novel	hum
482	26	83.9	571	7	ADC90359	Adc90359	Novel	hum
483	26	83.9	571	7	ADC69778	Adc69778	Human	PRO
484	26	83.9	571	7	ADC48667	Adc48667	Human	PRO
485	26	83.9	571	7	ADD10196	Add10196	Human	PRO
486	26	83.9	571	7	ADC78157	Adc78157	Novel	hum
487	26	83.9	571	7	ADD04771	Add04771	Novel	hum
488	26	83.9	571	7	ADD06392	Add06392	Novel	hum
489	26	83.9	571	7	ADC80727	Adc80727	Novel	hum
490	26	83.9	571	7	ADD11234	Add11234	Human	PRO
491	26	83.9	571	7	ADD10555	Add10555	Human	sec
492	26	83.9	571	7	ADC48115	Adc48115	Human	PRO
493	26	83.9	571	7	ADC77911	Adc77911	Novel	hum
494	26	83.9	571	7	ADC80175	Adc80175	Novel	hum
495	26	83.9	571	7	ADD11515	Add11515	Human	sec
496	26	83.9	571	7	ADD09644	Add09644	Human	PRO
497	26	83.9	571	7	ADD50874	Add50874	Novel	hum
498	26	83.9	571	7	ADD41357	Add41357	Novel	hum
499	26	83.9	571	7	ADD52496	Add52496	Human	PRO
500	26	83.9	571	7	ADD51120	Add51120	Novel	hum

ALIGNMENTS

RESULT 1	
ADQ90196	ADQ90196 standard; peptide; 6 AA.
ID	
AC	ADQ90196;
XX	
DT	21-OCT-2004 (first entry) *
XX	
DE	FSRYAR 6 mer domain motoneuronotrophic factor peptide analogue.
KW	motoneuronotrophic; motor neuron; Neurotrophic; Neuroprotective; Vulnerary;
KW	Neuropeptide-Agonist; axon degeneration; muscle reinnervation;
KW	peripheral nerve; neurodegenerative; wound healing.
OS	Unidentified.
XX	
PN	WO2004065410-A2.
XX	
PD	05-AUG-2004.
XX	
PE	21-JAN-2004; 2004WO-US001468.
XX	
PR	21-JAN-2003; 2003US-0441772P.
PA	(GENE-) GENERVON BIOPHARMACEUTICALS LLC.
XX	
PI	Chau RMW, Ko TD;
XX	
DR	WPI; 2004-562147/54.

XX New motoneuronotrophic factor peptide analogue enhancing the viability of
PT motor neurons, useful for target muscle reinnervation, treating
PT peripheral nerve injuries or neurodegenerative diseases and in wound
PT healing.
PS Claim 1; SEQ ID NO 2; 40pp; English.
XX
CC The present invention relates to motoneuronotrophic factor peptide
CC analogue, where the peptide analogue enhances the viability of motor
CC neurons. The methods and compositions of the present invention are useful
CC for promoting motor neuron viability and axon degeneration, target muscle
CC reinnervation, treating peripheral nerve injuries, treating
CC neurodegenerative disease and in wound healing. The present sequence
CC represents FSRYAR 6 mer domain motoneuronotrophic factor peptide
CC analogue.
SQ Sequence 6 AA;
Query Match 100.0%; Score 31; DB 8; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.1e+06; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;
OY 1 FSRYAR 6
DB 1 FSRYAR 6
RESULT 2
ADQ90198
ID ADQ90198 standard; peptide; 10 AA.
AC
XX ADQ90198;
XX
XX 21-OCT-2004 (first entry)
XX
XX MNTF 10 mer motoneuronotrophic factor peptide analogue.
XX
XX motoneuronotrophic; motor neuron; Neurotrophic; Neuroprotective; Vulnerary;
XX Neuropeptide-Agonist; axon degeneration; muscle reinnervation;
XX peripheral nerve; neurodegenerative; wound healing.
XX
XX Unidentified.
XX
XX WO2004065410-A2.
XX
XX 05-AUG-2004.
XX
XX 21-JAN-2004; 2004WO-US001468.
XX
XX 21-JAN-2003; 2003US-0441772P.
XX
XX (GENE-) GENERVON BIOPHARMACEUTICALS LLC.
XX
XX Chau RMW, Ko TD;
XX
XX WPI; 2004-562147/54.
XX
XX New motoneuronotrophic factor peptide analogue enhancing the viability of
PT motor neurons, useful for target muscle reinnervation, treating
PT peripheral nerve injuries or neurodegenerative diseases and in wound
PT healing.
PS Claim 7; SEQ ID NO 4; 40pp; English.
XX
XX The present invention relates to motoneuronotrophic factor peptide
XX analogue, where the peptide analogue enhances the viability of motor
XX neurons. The methods and compositions of the present invention are useful
XX for promoting motor neuron viability and axon degeneration, target muscle
XX reinnervation, treating peripheral nerve injuries, treating
XX neurodegenerative disease and in wound healing. The present sequence
XX represents a MNTF motoneuronotrophic factor peptide analogue.

SQ Sequence 10 AA;
Query Match 100.0%; Score 31; DB 8; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.2; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;
OY 1 FSRYAR 6
DB 5 FSRYAR 10
RESULT 3
ADQ90199
ID ADQ90199 standard; peptide; 11 AA.
AC
XX ADQ90199;
XX
XX 21-OCT-2004 (first entry)
XX
XX MNTF 11 mer motoneuronotrophic factor peptide analogue.
XX
XX motoneuronotrophic; motor neuron; Neurotrophic; Neuroprotective; Vulnerary;
XX Neuropeptide-Agonist; axon degeneration; muscle reinnervation;
XX peripheral nerve; neurodegenerative; wound healing.
XX
XX Unidentified.
XX
XX WO2004065410-A2.
XX
XX 05-AUG-2004.
XX
XX 21-JAN-2004; 2004WO-US001468.
XX
XX 21-JAN-2003; 2003US-0441772P.
XX
XX (GENE-) GENERVON BIOPHARMACEUTICALS LLC.
XX
XX Chau RMW, Ko TD;
XX
XX WPI; 2004-562147/54.
XX
XX New motoneuronotrophic factor peptide analogue enhancing the viability of
PT motor neurons, useful for target muscle reinnervation, treating
PT peripheral nerve injuries or neurodegenerative diseases and in wound
PT healing.
PS Claim 7; SEQ ID NO 5; 40pp; English.
XX
XX The present invention relates to motoneuronotrophic factor peptide
XX analogue, where the peptide analogue enhances the viability of motor
XX neurons. The methods and compositions of the present invention are useful
XX for promoting motor neuron viability and axon degeneration, target muscle
XX reinnervation, treating peripheral nerve injuries, treating
XX neurodegenerative disease and in wound healing. The present sequence
XX represents a MNTF motoneuronotrophic factor peptide analogue.
SQ Sequence 11 AA;
Query Match 100.0%; Score 31; DB 8; Length 11;
Best Local Similarity 100.0%; Pred. No. 3.6; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;
OY 1 FSRYAR 6
DB 1 FSRYAR 6
RESULT 4
ADQ90200
ID ADQ90200 standard; peptide; 13 AA.
AC
XX ADQ90200;
XX

DT 21-OCT-2004 (first entry)
 XX MNMF 13 mer motoneuronotrophic factor peptide analogue.
 DE
 XX motoneuronotrophic; motor neuron; Nootropic; Neuroprotective; Vulnerary;
 KW Neuropeptide-Agonist; axon degeneration; muscle reinnervation;
 KW peripheral nerve; neurodegenerative; wound healing.
 XX Unidentified.
 OS
 XX WO2004065410-A2.
 PN
 XX 05-AUG-2004.
 PD
 XX 21-JAN-2004; 2004WO-US001468.
 PF
 XX 21-JAN-2003; 2003US-0441772P.
 PR
 XX (GENE-) GENERVON BIOPHARMACEUTICALS LLC.
 PA
 XX Chau RMW, Ko TD;
 PI
 XX WPI; 2004-562147/54.
 DR
 XX New motoneuronotrophic factor peptide analogue enhancing the viability of
 PT motor neurons, useful for target muscle reinnervation, treating
 PT peripheral nerve injuries or neurodegenerative diseases and in wound
 PT healing.
 PS
 XX Claim 7; SEQ ID NO 6; 40pp; English.
 XX
 CC The present invention relates to motoneuronotrophic factor peptide
 CC analogue, where the peptide analogue enhances the viability of motor
 CC neurons. The methods and compositions of the present invention are useful
 CC for promoting motor neuron viability and axon degeneration, target muscle
 CC reinnervation, treating peripheral nerve injuries, treating
 CC neurodegenerative disease and in wound healing. The present sequence
 CC represents a MNMF motoneuronotrophic factor peptide analogue.
 XX
 SQ Sequence 13 AA;
 Query Match 100.0%; Score 31; DB 8; Length 13;
 Best Local Similarity 100.0%; Pred. No. 4.2;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FSRYAR 6
 Db 7 FSRYAR 12
 RESULT 5
 ADQ90201
 ID ADQ90201 standard; peptide; 21 AA.
 AC
 ADQ90201;
 DT 21-OCT-2004 (first entry)
 DE MNMF 21 mer motoneuronotrophic factor peptide analogue.
 DE
 KW motoneuronotrophic; motor neuron; Nootropic; Neuroprotective; Vulnerary;
 KW Neuropeptide-Agonist; axon degeneration; muscle reinnervation;
 KW peripheral nerve; neurodegenerative; wound healing.
 XX Unidentified.
 OS
 XX WO2004065410-A2.
 PN
 XX 05-AUG-2004.
 PD
 XX 21-JAN-2004; 2004WO-US001468.
 PR
 XX 21-JAN-2003; 2003US-0441772P.
 PR

XX (GENE-) GENERVON BIOPHARMACEUTICALS LLC.
 PA
 XX Chau RMW, Ko TD;
 PI
 XX WPI; 2004-562147/54.
 DR
 XX New motoneuronotrophic factor peptide analogue enhancing the viability of
 PT motor neurons, useful for target muscle reinnervation, treating
 PT peripheral nerve injuries or neurodegenerative diseases and in wound
 PT healing.
 PS
 XX Claim 7; SEQ ID NO 7; 40pp; English.
 XX
 CC The present invention relates to motoneuronotrophic factor peptide
 CC analogue, where the peptide analogue enhances the viability of motor
 CC neurons. The methods and compositions of the present invention are useful
 CC for promoting motor neuron viability and axon degeneration, target muscle
 CC reinnervation, treating peripheral nerve injuries, treating
 CC neurodegenerative disease and in wound healing. The present sequence
 CC represents a MNMF motoneuronotrophic factor peptide analogue.
 XX
 SQ Sequence 21 AA;
 Query Match 100.0%; Score 31; DB 8; Length 21;
 Best Local Similarity 100.0%; Pred. No. 6.9;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FSRYAR 6
 Db 5 FSRYAR 10
 RESULT 6
 AEA17675
 ID AEA17675 standard; protein; 21 AA.
 AC
 AEA17675;
 DT 28-JUL-2005 (first entry)
 DE Motoneuronotrophic factor (MNMF) polypeptide, SEQ ID NO:29.
 DE
 KW motoneuronotrophic factor; MNMF; cell growth; cell proliferation;
 KW DNA detection; chromosome 16.
 XX Homo sapiens.
 OS
 XX WO2005047487-A2.
 PN
 XX 26-MAY-2005.
 PD
 XX 08-NOV-2004; 2004WO-US038651.
 PR
 XX 07-NOV-2003; 2003US-0518581P.
 PA
 XX (GENE-) GENERVON BIOPHARMACEUTICALS LLC.
 PI
 XX Xue BB;
 XX
 XX WPI; 2005-386358/39.
 DR
 XX New isolated polynucleotide consisting of a motoneuronotrophic factor
 PT (MNMF)-associated nucleic acid sequence, useful for promoting the
 PT survival, growth, proliferation and maintenance of mammalian neurons.
 XX
 PS
 XX Claim 7; SEQ ID NO 29; 34pp; English.
 XX
 CC The invention relates to an isolated polynucleotide consisting of a
 CC motoneuronotrophic factor (MNMF)-associated nucleic acid sequence
 CC comprising AEA17647 or AEA17648, its fragment or complement. Also
 CC described are the following: a composition comprising a first and second
 CC polynucleotide, the first polynucleotide containing an open reading frame

CC ; an isolated MNTF associated polypeptide encoded by an open reading
CC frame of SEQ ID No. 1, a fusion protein comprising an MNTF associated
CC polypeptide encoded by an open reading frame of SEQ ID no. 1 linked to a
CC heterologous protein; an expression vector operably linked to the
CC isolated polynucleotide, where at least one open reading frame is
CC operably linked to a control sequence compatible with a desired host
CC vector; an isolated host cell transformed with the expression vector; a
CC method for determining the presence of a MNTF-associated polynucleotide
CC in a medium; a method of comparing the relative abundance of MNTF
CC associated expression products in different samples; and a panel for use
CC in hybridization assay comprising two or more polynucleotides stably
CC associated with the surface of a solid support. The polynucleotides are
CC useful for promoting the survival, growth, proliferation and maintenance
CC of mammalian neurons. The present sequence represents a polypeptide
CC encoded by an open reading frame of MNTF cDNA (see ABA17647), which is
CC located on chromosome 16q22.

XX
SQ Sequence 21 AA;

Query Match 100.0%; Score 31; DB 9; Length 21;
Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRYPAR 6
|||||
Db 5 FSRYPAR 10

RESULT 7
ID AAW59046 standard; protein; 33 AA.
XX AAW59046;
XX
DT 11-AUG-1998 (first entry)
XX
DE Human MNTF1-F6 protein fragment.
XX
XX Motoneuronotrophic factor; MNTF-1; MNTF1-F6; human; axon regeneration;
KM motoneuron; diagnose; treatment; disease; wound healing; scar tissue;
KM keloid.
XX
OS Homo sapiens.
XX
XX WO9813492-A2.
XX
XX 02-APR-1998.
XX
XX 22-SEP-1997; 97WO-US017142.
XX
XX 27-SEP-1996; 96US-0026792P.
XX 15-NOV-1996; 96US-00751225.
XX 12-SEP-1997; 97US-00928862.
XX
XX (KMBI-) KM BIOTECH INC.
XX
XX Chau RMW;
XX
XX MPI: 1998-230703/20.
XX
XX N-PSDB; AAV11748.
XX
XX Motoneuronotrophic factor MNTF1-F3 and MNTF1-F6 - useful for motoneuron
XX regeneration, diagnosing or treating motoneuron disease and to accelerate
XX wound healing without scar formation.
XX
XX PS Claim 4; Fig 2B; 78pp; English.
XX
XX This sequence, represents a fragment of a novel human motoneuronotrophic
XX factor, MNTF1-F6. Such factors are used to promote regeneration of the
XX axon of a motoneuron, to diagnose and treat motoneuron disease in a
XX mammal or to accelerate wound healing whilst concomitantly minimising or
XX inhibiting scar tissue and/or keloid formation in an area associated with
XX a wound. For promoting axonal regeneration, the polypeptide is

CC administered at a concentration of 5 ng-50 mg, whereas for inhibiting
CC hereditary motoneuron disease, the dosage is 5-100 (especially 30-50)ng
CC per kg body weight
XX
XX Sequence 33 AA;

QY 1 FSRYPAR 6
|||||
Db 17 FSRYPAR 22

RESULT 8
ID AAO29914 standard; protein; 33 AA.
XX AAO29914;
XX
AC AAO29914;
XX
DT 03-SEP-2003 (first entry)
XX
DE Human motoneuronotrophic factor (MNTF)1-F6 protein.
XX
XX Human; motoneuronotrophic factor; MNTF; therapy; radiation therapy;
KM peripheral nerve injury; musculoskeletal disorder; spinal cord injury;
KM head injury; stroke; neuromuscular degenerative disease; noctropic;
KM amyotrophic lateral sclerosis; peripheral neuropathy; neuroprotective;
KM muscular dystrophy; AIDS; spinal muscular atrophy; multiple sclerosis;
KM scar; myasthenia gravis; sensory neuronal function disorder; vulnerability;
KM cancer.
XX
XX Homo sapiens.
XX
XX WO2003044175-A2.
XX
XX 30-MAY-2003.
XX
XX 19-NOV-2002; 2002WO-US037191.
XX
XX 20-NOV-2001; 2001US-00989481.
XX
XX (GENE-) GENERON BIOPHARMACEUTICALS LLC.
XX
XX Chau RMW;
XX
XX MPI: 2003-457607/43.
XX
XX N-PSDB; AAL60573.
XX
XX Promoting the survival, growth, proliferation or maintenance of mammalian
XX neurons by administering motoneuronotrophic factors, useful for treating
XX musculoskeletal and neurodegenerative disorders and spinal cord injuries.
XX
XX Claim 1; Fig 2B; 90pp; English.
XX
XX The invention relates to a method for promoting the survival, growth,
XX proliferation or maintenance of mammalian neurons by administering
XX motoneuronotrophic factors (MNTF). The method is useful for treating
XX peripheral nerve injuries, musculoskeletal disorders, spinal cord
XX injuries, head injuries, strokes, neuromuscular degenerative diseases,
XX amyotrophic lateral sclerosis, spinal muscular atrophy, peripheral
XX neuropathy, diabetic peripheral neuropathy, peripheral neuropathy
XX resulting from AIDS or radiation therapy for cancer, multiple sclerosis,
XX muscular dystrophy, inhibition of scar formation, myasthenia gravis and
XX sensory neuronal function disorders. The present sequence is human MNTF1-
XX F6 protein used to illustrate the method of the invention

XX
SQ Sequence 33 AA;

Query Match 100.0%; Score 31; DB 6; Length 33;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 FSRYAR 6
 |||||
 Db 17 FSRYAR 22

RESULT 9
 ADQ90195

ID ADQ90195 standard; peptide; 33 AA.

XX AC ADQ90195;

XX DT 21-OCT-2004 (first entry)

XX DE MNTF1-F6 33 mer motoneuronotrophic factor peptide analogue.

XX KM motoneuronotrophic; motor neuron; Nootropic; Neuroprotective; Vulnerary;
 KM Neuropeptide-Agonist; axon degeneration; muscle reinnervation;
 KM peripheral nerve; neurodegenerative; wound healing.

XX OS Unidentified.

XX PN WO2004065410-A2.

XX PD 05-AUG-2004.

XX PF 21-JAN-2004; 2004WO-US001468.

XX PR 21-JAN-2003; 2003US-0441772P.

XX PA (GENE-) GENERVON BIOPHARMACEUTICALS LLC.

XX PI Chau RMM, Ko TD;

XX XX MPI, 2004-562147/54.

XX DR
 PT New motoneuronotrophic factor peptide analogue enhancing the viability of
 PT motor neurons, useful for target muscle reinnervation, treating
 PT peripheral nerve injuries or neurodegenerative diseases and in wound
 PT healing.

XX PS Claim 2; SEQ ID NO 1; 40pp; English.

XX CC The present invention relates to motoneuronotrophic factor peptide
 CC analogue, where the peptide analogue enhances the viability of motor
 CC neurons. The methods and compositions of the present invention are useful
 CC for promoting motor neuron viability and axon degeneration, target muscle
 CC reinnervation, treating peripheral nerve injuries, treating muscle
 CC neurodegenerative disease and in wound healing. The present sequence
 CC represents MNTF1-F6 33 mer motoneuronotrophic factor peptide analogue.

XX SQ Sequence 33 AA;

Query Match 100.0%; Score 31; DB 8; Length 33;

Best Local Similarity 100.0%; Pred. No. 11;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 FSRYAR 6
 |||||
 Db 17 FSRYAR 22

RESULT 10

ID AAU57756 standard; protein; 81 AA.

XX AC AAU57756;

XX DT 27-FEB-2002 (first entry)

XX DE Propionibacterium acnes immunogenic protein #18652.

XX KM SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;

KM uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KM inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KM dermatological; osteopathic; neuroprotectant.

XX OS Propionibacterium acnes.

XX PN WO200181581-A2.

XX PD 01-NOV-2001.

XX PE 20-APR-2001; 2001WO-US012865.

XX PR 21-APR-2000; 2000US-0199047P.

XX PR 02-JUN-2000; 2000US-0208841P.

XX PR 07-JUL-2000; 2000US-0216747P.

XX PA (CORI-) CORIXA CORP.

XX PI Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX DR MPI; 2001-616774/71.

XX DR N-PSDB; AAS59585.

XX PT Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris.

XX PS Example 1; SEQ ID NO 18951; 1069pp; English.

XX CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pcl_sequences

XX SQ Sequence 81 AA;

Query Match 100.0%; Score 31; DB 4; Length 81;

Best Local Similarity 100.0%; Pred. No. 28;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 FSRYAR 6
 |||||
 Db 14 FSRYAR 19

RESULT 11

ID ABM54275 standard; protein; 81 AA.

XX AC ABM54275;

XX DT 20-OCT-2003 (first entry)

XX DE Propionibacterium acnes predicted ORF-encoded polypeptide #18951.

XX KM Acne vulgaris; antiseborrhoeic; dermatological; antibacterial;

XX	IMUNOSTIMULANT; immune response; vaccine.
XX	Propionibacterium acnes.
OS	WO2003033515-A1.
XX	24-APR-2003.
XX	11-OCT-2002; 2002WO-US032727.
XX	15-OCT-2001; 2001US-00978825.
XX	(CORI-) CORIXA CORP.
XX	Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL, Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D, Barth B, Valliave-Dougllass J;
XX	WPI; 2003-381789/36.
XX	N-PSDB; ACP64514.
XX	New Propionibacterium acnes polypeptides and polynucleotides encoding the polypeptide, useful for diagnosing, preventing or treating acne vulgaris, or for stimulating an immune response specific for a P. acnes protein.
XX	Example 1; SEQ ID NO 18951; 1481bp; English.
XX	The invention relates to an isolated polynucleotide (ACF64435-ACF64733) encoding a Propionibacterium acnes protein. The invention also relates to polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to immunogenic fragments of P. acnes polypeptides. The invention additionally encompasses expression vectors and host cells comprising a polynucleotide of the invention; antibodies against polypeptides of the invention; fusion proteins comprising a polypeptide of the invention; a method for stimulating an immune response specific for a P. acnes polypeptide and an isolated T cell population comprising T cells prepared via this method; a vaccine composition (comprising P. acnes polypeptides, polynucleotides, antibodies, fusion proteins, T cell populations, or antigen-presenting cells that express the polypeptide); a method and kit for detecting or determining the presence or absence of P. acnes in a patient; and a method for inhibiting the development of P. acnes in a patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion proteins, T cell populations or antigen-presenting cells that express the polypeptides are useful for diagnosing, preventing or treating acne vulgaris, or for stimulating an immune response specific for a P. acnes protein. The polynucleotides can also be used as probes or primers for nucleic acid hybridisation. The vaccine composition is useful for the stimulation of an immune response against P. acnes, or for treating acne, and the kit is useful for performing a diagnostic assay. The present sequence represents a polypeptide predicted to be encoded by an ORF (open reading frame) contained within the P. acnes polynucleotides of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX	Sequence 81 AA:
XX	Query Match 100.0%; Score 31; DB 6; Length 81;
XX	Best Local Similarity 100.0%; Fred. No. 28;
XX	Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0
XX	1 FSRVAR 6
XX	14 FSRVAR 19
XX	RESULT 12
XX	AAU57258
XX	ID AAU57258 standard; protein; 112 AA.
XX	AAU57258;
XX	27-FEB-2002 (first entry)

Query Match	Best Local Similarity	Score 31;	DB 4;	Length 112;
Matches 6;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
OY 1 FSRVAR 6				
DB 35 FSRVAR 40				
RESULT 13				
ABM53777				
ID ABM53777	standard;	protein;	112 AA.	
XX ABM53777;				
XX				
DT 20-OCT-2003	(first entry)			

XX DE Propionibacterium acnes predicted ORF-encoded polypeptide #18453.
 XX XX Acne vulgaris; antiseborrheic; dermatological; antibacterial;
 KM immunostimulant; immune response; vaccine.
 XX OS Propionibacterium acnes.
 XX PN WO2003033515-A1.
 XX PD 24-APR-2003.
 XX PF 11-OCT-2002; 2002WO-US032727.
 XX PR 15-OCT-2001; 2001US-00978825.
 XX PA (CORI-) CORIXA CORP.
 XX PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
 PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
 PI Barth B, Vallieve-Douglas J;
 XX WPI: 2003-381789/36.
 DR N-PSDB: ACF64510.
 PT New Propionibacterium acnes polypeptides and polynucleotides encoding the
 PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
 PT or for stimulating an immune response specific for a P. acnes protein.
 XX Example 1; SEQ ID NO 18453; 1481bp; English.
 XX PS The invention relates to an isolated polynucleotide (ACF6435-ACF64733)
 CC encoding a Propionibacterium acnes protein. The invention also relates to
 CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
 CC immunogenic fragments of P. acnes polypeptides. The invention
 CC additionally encompasses expression vectors and host cells comprising a
 CC polynucleotide of the invention; antibodies against polypeptides of the
 CC invention; fusion proteins comprising a polypeptide of the invention; a
 CC method for stimulating an immune response specific for a P. acnes
 CC polypeptide and an isolated T cell population comprising T cells prepared
 CC via this method; a vaccine composition (comprising P. acnes polypeptides,
 CC polynucleotides, antibodies, fusion proteins, T cell populations, or
 CC antigen-presenting cells that express the polypeptide); a method and kit
 CC for detecting or determining the presence or absence of P. acnes in a
 CC patient; and a method for inhibiting the development of P. acnes in a
 CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
 CC proteins, T cell populations or antigen-presenting cells that express the
 CC polypeptides are useful for diagnosing, preventing or treating acne
 CC vulgaris, or for stimulating an immune response specific for a P. acnes
 CC protein. The polynucleotides can also be used as probes or primers for
 CC nucleic acid hybridisation. The vaccine composition is useful for the
 CC stimulation of an immune response against P. acnes, or for treating acne,
 CC and the kit is useful for performing a diagnostic assay. The present
 CC sequence represents a polypeptide predicted to be encoded by an ORF (open
 CC reading frame) contained within the P. acnes polynucleotide of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX SQ Sequence 112 AA;
 Query Match 100.0%; Score 31; DB 6; Length 112;
 Best Local Similarity 100.0%; Pred. No. 39;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FSRYAR 6
 Db 35 FSRYAR 40

RESULT 14
 ABM64633
 ID ABM64633 standard; protein; 174 AA.

XX AC ABM64633;
 XX XX 20-OCT-2003 (first entry)
 DT DE Propionibacterium acnes immunogenic polypeptide #29309.
 XX XX Acne vulgaris; antiseborrheic; dermatological; antibacterial;
 KM immunostimulant; immune response; vaccine; immunogenic.
 XX OS Propionibacterium acnes.
 XX PN WO2003033515-A1.
 XX PD 24-APR-2003.
 XX PF 11-OCT-2002; 2002WO-US032727.
 XX PR 15-OCT-2001; 2001US-00978825.
 XX PA (CORI-) CORIXA CORP.
 XX PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
 PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
 PI Barth B, Vallieve-Douglas J;
 XX WPI: 2003-381789/36.
 DR New Propionibacterium acnes polypeptides and polynucleotides encoding the
 DR polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
 DR or for stimulating an immune response specific for a P. acnes protein.
 XX Claim 7; SEQ ID NO 29309; 1481bp; English.
 XX PS The invention relates to an isolated polynucleotide (ACF6435-ACF64733)
 CC encoding a Propionibacterium acnes protein. The invention also relates to
 CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
 CC immunogenic fragments of P. acnes polypeptides. The invention
 CC additionally encompasses expression vectors and host cells comprising a
 CC polynucleotide of the invention; antibodies against polypeptides of the
 CC invention; fusion proteins comprising a polypeptide of the invention; a
 CC method for stimulating an immune response specific for a P. acnes
 CC polypeptide and an isolated T cell population comprising T cells prepared
 CC via this method; a vaccine composition (comprising P. acnes polypeptides,
 CC polynucleotides, antibodies, fusion proteins, T cell populations, or
 CC antigen-presenting cells that express the polypeptide); a method and kit
 CC for detecting or determining the presence or absence of P. acnes in a
 CC patient; and a method for inhibiting the development of P. acnes in a
 CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
 CC proteins, T cell populations or antigen-presenting cells that express the
 CC polypeptides are useful for diagnosing, preventing or treating acne
 CC vulgaris, or for stimulating an immune response specific for a P. acnes
 CC protein. The polynucleotides can also be used as probes or primers for
 CC nucleic acid hybridisation. The vaccine composition is useful for the
 CC stimulation of an immune response against P. acnes, or for treating acne,
 CC and the kit is useful for performing a diagnostic assay. The present
 CC sequence represents a specifically claimed P. acnes polypeptide which is
 CC thought to contain an immunogenic region. Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX SQ Sequence 174 AA;
 Query Match 100.0%; Score 31; DB 6; Length 174;
 Best Local Similarity 100.0%; Pred. No. 61;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FSRYAR 6
 Db 35 FSRYAR 40

```

RESULT 15
AAV35559
ID AAV35559 standard; protein; 357 AA.
XX
XX AAV35559;
AC
XX 17-OCT-2003 (revised)
DT 13-SEP-1999 (first entry)
XX
DE Chlamydia pneumoniae transmembrane protein sequence.
XX
XX Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
KM sinusitis; purulent otitis media; erythema nodosum; pharyngitis; vaccine;
KW neutralising epitope.
XX
XX Chlamydia pneumoniae.
OS
XX WO9927105-A2.
XX
XX 03-JUN-1999.
XX
XX 20-NOV-1998; 98WO-IB001890.
XX
XX 21-NOV-1997; 97FR-00014673.
PR 04-NOV-1998; 98US-0107078P.
XX
XX (GEST ) GENSET.
XX
XX Griflais R;
PI
XX WPI; 1999-357842/30.
XX
XX Genome sequence of Chlamydia pneumoniae.
PT
XX Page 1303-1304; Disclosure; 1912pp; English.
XX
XX AAV34584-Y35879 represent the proteins encoded by all the open reading
CC frames in the complete genome (see AAX91990) of Chlamydia pneumoniae. C.
CC pneumoniae causes respiratory disease such as pneumonia and bronchitis
CC and is thought to be a contributing factor in heart disease, sarcoidosis,
CC sinusitis, purulent otitis media, erythema nodosum or pharyngitis. The
CC polypeptides encoded by the open reading frames of the C. pneumoniae
CC genome (see AAV34584-Y35879) can be used in immunogenic compositions as
CC vaccines. Vectors containing C. pneumoniae nucleotide sequences can also
CC be used as immunogenic compositions, especially where the vector directs
CC the expression of a neutralising epitope of C. pneumoniae. (Updated on 17
CC -OCT-2003 to standardise OS field)
XX
XX SQ Sequence 357 AA;
XX
XX Query Match 100.0%; Score 31; DB 2; Length 357;
XX Best Local Similarity 100.0%; Pred. No. 1.3e+02;
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 FSRYAR 6
XX |||||
DB 135 FSRYAR 140
XX

RESULT 16
ABB90579
ID ABB90579 standard; protein; 357 AA.
XX
XX ABB90579;
AC
XX 29-AUG-2003 (revised)
DT 29-JUL-2002 (first entry)
XX
XX Chlamydia pneumoniae CP7224 protein, SEQ ID NO:107.
DE
XX Chlamydia infection; antigen; immunogen; vaccine; diagnosis;
KM human respiratory disease; cardiovascular disease; atherosclerosis;
KW coronary artery disease; carotid artery stenosis; myocardial infarction;
XX

```

```

KM cerebrovascular disease; aortic aneurysm; claudication; stroke;
KM strain CWL029.
XX
XX Chlamydia pneumoniae.
OS
XX WO200202606-A2.
XX
XX 10-JAN-2002.
XX
XX 03-JUL-2001; 2001WO-IB001445.
XX
XX 03-JUL-2000; 2000GB-00016363.
PR 11-JUL-2000; 2000GB-00017047.
PR 21-JUL-2000; 2000GB-00017983.
PR 07-AUG-2000; 2000GB-00019368.
PR 18-AUG-2000; 2000GB-00020440.
PR 14-SEP-2000; 2000GB-00022583.
PR 10-NOV-2000; 2000GB-00027549.
PR 22-DEC-2000; 2000GB-00031706.
XX
XX (CHIR-) CHIRON SPA.
XX
XX Ratti G, Grandi G;
PI
XX WPI; 2002-154726/20.
XX
XX N-PSDB; ABL91237.
XX
XX Novel Chlamydia pneumoniae protein useful in the manufacture of a
PT medicament for treatment or prevention of infection due to Chlamydia,
PT preferably Chlamydia pneumoniae, and for diagnostic purposes.
XX
XX Claim 1; Page 93; 364pp; English.
XX
XX Sequences ABB90526-ABB90715 represent novel proteins from Chlamydia
CC pneumoniae (strain CWL029), and ABL9184-ABL91373 represent DNA encoding
CC them. The proteins are predicted to be immunogenic and may therefore be
CC useful in vaccine production and for diagnostic purposes. Chlamydia
CC pneumoniae is a common cause of respiratory disease in humans, and is
CC also involved in the development of cardiovascular diseases such as
CC atherosclerosis, coronary artery disease, carotid artery stenosis,
CC myocardial infarction, cerebrovascular disease, aortic aneurysm,
CC claudication and stroke. The proteins and nucleic acids of the invention
CC may be used in vaccines and pharmaceutical compositions for the
CC prevention or treatment of chlamydial infections, particularly Chlamydia
CC pneumoniae infections. The proteins may also be used in the detection of
CC Chlamydia pneumoniae, and the nucleic acids may be used in PCR, branched
CC DNA probe assay or blotting techniques for determining Chlamydia
CC pneumoniae gene expression. The present sequence represents a
CC specifically claimed Chlamydia pneumoniae protein of the invention.
CC (Updated on 29-AUG-2003 to standardise OS field)
XX
XX SQ Sequence 357 AA;
XX
XX Query Match 100.0%; Score 31; DB 5; Length 357;
XX Best Local Similarity 100.0%; Pred. No. 1.3e+02;
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 FSRYAR 6
XX |||||
DB 135 FSRYAR 140
XX

RESULT 17
ABU27032
ID ABU27032 standard; protein; 357 AA.
XX
XX ABU27032;
AC
XX 23-OCT-2003 (revised)
DT 19-JUN-2003 (first entry)
XX
XX Protein encoded by Prokaryotic essential gene #12559.
DE
XX

```

KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX Chlamydia pneumoniae.
 OS
 PN WO200277183-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH,
 XX
 DR WPI; 2003-029926/02.
 DR N-PSDB; ACAA30902.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 54956; 1766pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-regulated gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC parent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_ptc_sequences. (Updated on 23-OCT-2003 to
 CC standardise OS field)
 XX
 SQ Sequence 357 AA;
 Query Match 100.0%; Score 31; DB 6; Length 357;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FSRYAR 6
 |||||
 DB 135 FSRYAR 140

RESULT 18
 AEC95703
 ID AEC95703 standard; protein; 357 AA.
 XX
 AC AEC95703;
 XX
 DT 17-NOV-2005 (first entry)
 XX
 DE C. pneumoniae Murg peptidoglycan transferase protein (Cpn 0904).
 XX
 KW diagnosis; autotransporter antigen; antibacterial; vaccine;
 KW Murg peptidoglycan transferase protein; Cpn 0904.
 XX
 OS Chlamydia pneumoniae.
 XX
 PN WO2005084306-A2.
 XX
 PD 15-SEP-2005.
 XX
 PF 02-MAR-2005; 2005WO-US006588.
 XX
 PR 02-MAR-2004; 2004US-0549832P.
 PR 12-JAN-2005; 2005US-0643110P.
 PR 19-JAN-2005; 2005US-0644552P.
 XX
 PA (CHIR) CHIRON CORP.
 XX
 PI Grandi G, Giulio R;
 XX
 DR WPI; 2005-619256/63.
 DR GENBANK; AAD19042.1.
 XX
 PT New Chlamydia pneumoniae autotransporter protein, useful as an
 PT autotransporter antigen, for raising a specific immune response or for
 PT preventing or treating C. pneumoniae infections in individuals.
 XX
 PS Disclosure; SEQ ID NO 18; 159pp; English.
 XX
 CC The invention relates to a polypeptide for use as an autotransporter
 CC antigen. Also provided is a method of eliciting an immune response in an
 CC individual which involves administering to the individual an Chlamydia
 CC pneumoniae autotransporter protein selected from Cpn0794, Cpn0795,
 CC Cpn0796, Cpn0797, Cpn0798, or Cpn0799 or an immunogenic fragment. The
 CC polypeptide is secreted into the cytoplasm of the host cell through a
 CC Type V autotransporter secretion system mechanism. The polypeptide is
 CC useful as an antigen for raising a C. pneumoniae specific immune
 CC response. It is also useful for raising a systemic immune response in an
 CC individual infected with C. pneumoniae. The polypeptide is also useful in
 CC diagnosis. It can also be used in the preparation of a medicament for the
 CC prevention or treatment of a C. pneumoniae infection in an individual, or
 CC in the preparation of an assay for the diagnosis of a C. pneumoniae
 CC infection in an individual. It can also be used as an autotransporter
 CC protein which immunoreacts with seropositive serum of an individual
 CC infected with C. pneumoniae. The present sequence represents a C.
 CC pneumoniae Murg peptidoglycan transferase protein (Cpn 0904).
 XX
 SQ Sequence 357 AA;
 Query Match 100.0%; Score 31; DB 9; Length 357;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FSRYAR 6
 |||||
 DB 135 FSRYAR 140
 RESULT 19
 ADS44828
 ID ADS44828 standard; protein; 463 AA.
 XX
 AC ADS44828;

XX 02-DEC-2004 (first entry)
XX
XX
XX Bacterial polypeptide #23258.
XX
XX Recombinant DNA construct; transformed plant; improved plant property;
XX cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
XX pathogen tolerance; pest tolerance; plant disease resistance;
XX cell cycle pathway modification; plant growth regulator;
XX homologous recombination; seed oil yield; protein yield; carbohydrate;
XX nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
XX bacterial polypeptide.
XX
XX Bacteria.
XX
XX US2003233675-A1.
XX
XX
XX PD 18-DEC-2003.
XX
XX PF 20-FEB-2003; 2003US-00369493.
XX
XX PR 21-FEB-2002; 2002US-0360039P.
XX
XX PA (CAOY/ CAO Y.
XX (HINK/ HINKLE G J.
XX (SLAT/ SLATER S C.
XX (CHEN/ CHEN X.
XX (GOLD/ GOLDMAN B S.
XX
XX PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX
XX WPI, 2004-061375/06.
XX
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
XX for expression of a polynucleotide encoding a polypeptide from a
XX microbial source, useful for producing plants with improved properties.
XX
XX
XX Claim 1; SEQ ID NO 23258; 122bp; English.
XX
XX
XX The invention relates to a recombinant DNA construct comprising a
XX promoter functional in a plant cell, where the promoter is positioned to
XX provide for expression of a polynucleotide encoding a polypeptide from a
XX microbial source. The invention also relates to a transformed plant
XX comprising the recombinant DNA construct and a method of producing a
XX such as maize or soybean. The method of producing a transformed plant
XX having an improved property comprises transforming a plant with the
XX recombinant DNA construct and growing the transformed plant, where the
XX polynucleotide or polypeptide is useful for improving plant properties.
XX The recombinant DNA construct is useful for producing plants with
XX improved plant properties, e.g. improved cold, heat or drought tolerance,
XX tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
XX increased resistance to plant disease, better growth rate by modification
XX of the cell cycle pathway with plant growth regulators, increased rate of
XX homologous recombination, modified seed oil or protein yield and/or
XX content, improved yield by modification of carbohydrate, nitrogen or
XX phosphorus use and/or uptake, by modification of photosynthesis or by
XX providing improved plant growth and development under at least one stress
XX condition, improved lignin production or improved galactomannan
XX production. This sequence represents a bacterial polypeptide used in the
XX scope of the invention. Note: The sequence data for this patent did not
XX form part of the printed specification but was obtained in electronic
XX format from USPTO at seqdata.uspto.gov/sequence.html.
XX
XX Sequence 463 AA:
XX

Query Match	100.0%;	Score 31;	DB 8;	Length 463;
Best Local Similarity	100.0%;	Pred. No. 1.7e+02;		
Matches 6;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	FSRYAR	6	
Db	76	FSRYAR	81	

RESULT 20	
AAE23630	
ID	AAE23630 standard; protein; 466 AA.
XX	
AC	AAE23630;
XX	
DT	27-AUG-2002 (first entry)
XX	
DE	Escherichia coli 6-phospho-strand-glucosidase #1.
XX	
KW	Secondary structural element: genome-sized database; EC 3.2.1.86; enzyme: 6-phospho-strand-glucosidase.
XX	
OS	Escherichia coli.
XX	
PH	Key
FT	Location/Qualifiers
FT	9..11
FT	/note="Strand A"
FT	12..103
FT	/note="GG dipeptide parse"
FT	14..20
FT	/note="Strand B"
FT	48..59
FT	/note="Parse region"
FT	49..51
FT	/note="Strand A"
FT	50..54
FT	/note="Parsing string (PGDSG)"
FT	53..60
FT	/note="Strand B"
FT	62..63
FT	/note="GG dipeptide parse"
FT	71..81
FT	/note="Parse region"
FT	72..84
FT	/note="Helix 1"
FT	73..78
FT	/note="Beta strand"
FT	89..93
FT	/note="Strand C"
FT	95..107
FT	/note="Helix 1"
FT	95..102
FT	/note="Active site a"
FT	111..115
FT	/note="Strand C"
FT	116..130
FT	/note="Helix 2"
FT	117..125
FT	/note="Active site a"
FT	136..140
FT	/note="Strand D"
FT	138..153
FT	/note="Helix 2"
FT	141
FT	/note="Active site b"
FT	158..177
FT	/note="Helix 3"
FT	159..163
FT	/note="Strand D"
FT	164..166
FT	/note="Active site b"
FT	181..198
FT	/note="Helix 3"
FT	182..185
FT	/note="Strand E"
FT	184..187
FT	/note="Active site c"
FT	205..209
FT	/note="Active site c"
FT	205..208
FT	Region

```

FT      /note= "Strand E"
FT      207..209
FT      /note= "Active site c"
FT      212..226
FT      /note= "Helix 4"
FT      212..215
FT      /note= "Beta strand"
FT      219..221
FT      /note= "Beta strand"
FT      242..245
FT      /note= "Strand X"
FT      248..268
FT      /note= "Helix 4"
FT      259..273
FT      /note= "Helix X"
FT      275..280
FT      /note= "Strand Y"
FT      286..293
FT      /note= "Helix 5"
FT      318..320
FT      /note= "Strand Y"
FT      323..327
FT      /note= "Strand F"
FT      329..339
FT      /note= "Helix Y"
FT      332..342
FT      /note= "Helix 5"
FT      375..382
FT      /note= "Strand Z"
FT      381..388
FT      /note= "Strand F"
FT      385..398
FT      /note= "Helix 6"
FT      404..407
FT      /note= "Strand G"
FT      408..410
FT      /note= "Active site d"
FT      431..448
FT      /note= "Helix 7"
FT      446..452
FT      /note= "Strand Z"
FT      450..454
FT      /note= "Strand H"
FT      456..469
FT      /note= "Helix 6"
FT      456..459
FT      /note= "Strand I"
FT      460..462
FT      /note= "Dipeptide GP parse"
FT      464..467
FT      /note= "Strand J"
FT      476..479
FT      /note= "Strand G"
FT      478..482
FT      /note= "Strand K"
FT      480..482
FT      /note= "Active site d"
FT      496..509
FT      /note= "Helix 8"
FT      497..517
FT      /note= "Helix 7"
FT      521..525
FT      /note= "Strand H"
FT      527..530
FT      /note= "Strand I"
FT      535..539
FT      /note= "Strand J"
FT      548..554
FT      /note= "Strand K"
FT      563..576
FT      /note= "Helix 8"
FT      XX
FN      US6377893-B1.

```

```

XX      23-APR-2002.
PD      19-AUG-1997; 97US-00914375.
XX      25-MAR-1992; 92US-00857224.
XX      (BENN/) BENNER S A.
XX      Benner SA;
XX      WPI; 2002-424771/45.
XX      Methods for excluding or detecting homology between protein families,
XX      useful e.g. for identifying in vitro properties of proteins important for
XX      physiological activity.
XX      Example 5; Col 147-150; 99pp; English.
XX      The invention relates to a method for excluding homology between two
XX      protein families. The method involves constructing models for secondary
XX      structural elements for each family; aligning secondary structural
XX      elements of one family with the secondary structural elements from the
XX      other family around sequence motifs; determining whether secondary
XX      structural elements flanking the sequence motifs in one family are
XX      congruent to secondary structural elements in the other family, so as to
XX      determine if the families are related by common ancestry or not. The
XX      method is used to confirm/deny the hypothesis that proteins are
XX      homologous and related methods are used to identify mutations during
XX      divergent evolution of proteins, to identify in vitro properties of
XX      proteins that are important for physiological activity and to generate
XX      genome-sized databases. The present sequence is Escherichia coli 6-
XX      phospho-strand-glucosidase (EC 3.2.1.86). This sequence is used in the
XX      exemplification of the invention
XX      Sequence 466 AA:
XX      SQ
XX      Query Match 100.0%; Score 31; DB 5; Length 466;
XX      Best Local Similarity 100.0%; Pred. No. 1.7e+02;
XX      Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY      1 FSRYAR 6
Db      154 FSRYAR 159
RESULT 21
AAU34677
ID      AUJ34677 standard; procein; 474 AA.
XX      AAU34677;
AC      14-FEB-2002 (first entry)
XX      DE      E. coli cellular proliferation protein #258.
XX      KW      Antisense; prokaryotic cellular proliferation protein; antibiotic;
XX      antibiotic; drug design.
XX      OS      Escherichia coli.
XX      PN      WO200170955-A2.
XX      PD      27-SEP-2001.
XX      PF      21-MAR-2001; 2001WO-US009180.
XX      PR      21-MAR-2000; 2000US-0191078P.
XX      PR      23-MAY-2000; 2000US-0206848P.
XX      PR      26-MAY-2000; 2000US-0207727P.
XX      PR      23-OCT-2000; 2000US-0242578P.
XX      PR      27-NOV-2000; 2000US-0253625P.
XX      PR      22-DEC-2000; 2000US-0257931P.

```

PR 16-FEB-2001; 2001US-0269308P.
XX
XX (ELITR-) ELITRA PHARM INC.
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ,
PI Yamamoto RT, Xu HH,
XX WPI; 2001-611495/70.
DR N-PSDB; AAS52536.
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids.
XX
PS Example 3; SEQ ID NO 10270; 511pp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC their use in the discovery of novel antibiotics, the essential genes
CC themselves and the encoded proteins. The prokaryotes used are *Escherichia*
CC *coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,
CC *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence represents an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences)
XX
SQ Sequence 474 AA;
Query Match 100.0%; Score 31; DB 4; Length 474;
Best Local Similarity 100.0%; Pred. No. 1.7e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;
QY 1 FSRYYR 6
Db 157 FSRYYR 162
RESULT 22
ADS45088
ID ADS45088 standard; protein; 474 AA.
XX
AC ADS45088;
XX
XX 02-DEC-2004 (first entry)
XX
DE Bacterial polypeptide #23518.
XX
XX Recombinant DNA construct; transformed plant; improved plant property;
KM cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KM pathogen tolerance; pest tolerance; plant disease resistance;
KM cell cycle pathway modification; plant growth regulator;
KM homologous recombination; seed oil yield; protein yield; carbohydrate;
KM nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KM bacterial polypeptide.
XX
XX Bacteria.
XX
XX US2003233675-A1.
XX
XX 18-DEC-2003.
XX
XX 20-FEB-2003; 2003US-00369493.
XX
XX 21-FEB-2002; 2002US-0360039P.
XX

PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
PI WPI; 2004-061375/06.
XX
DR WPI; 2004-061375/06.
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
PS Claim 1; SEQ ID NO 23518; 122pp; English.
XX
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 474 AA;
Query Match 100.0%; Score 31; DB 8; Length 474;
Best Local Similarity 100.0%; Pred. No. 1.7e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;
QY 1 FSRYYR 6
Db 157 FSRYYR 162
RESULT 23
ABO74595
ID ABO74595 standard; protein; 509 AA.
XX
XX ABO74595;
XX
XX 29-JUL-2004 (first entry)
XX
DE *Pseudomonas aeruginosa* polypeptide #6770.
XX
XX *Pseudomonas aeruginosa* polypeptide #6770.
XX
XX Bacterial infection; *Pseudomonas aeruginosa* infection; antibacterial.
XX
XX *Pseudomonas aeruginosa*.
XX
XX US6551795-B1.
XX
XX 22-APR-2003.
XX
XX 18-FEB-1999; 99US-00252991.
XX
XX 18-FEB-1998; 98US-0074788P.
XX

PR 27-JUL-1998; 98US-0094190P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX WPI; 2003-615309/58.
DR N-PSDB; ABD08166.
XX
PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
PS
PS Disclosure; SEQ ID NO 23341; 455pp; English.
XX
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABO67826-
CC ABO84336 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
SQ Sequence 509 AA;
OY
OY 1 FSRYYAR 6
OY |||||
Db 270 FSRYYAR 275
Query Match 100.0%; Score 31; DB 7; Length 509;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
RESULT 24
ABG24444
ID ABG24444 standard; protein; 1225 AA.
XX
AC ABG24444;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #24435.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS88631.
PT

XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
XX Claim 20; SEQ ID NO 54803; 103pp; English.
PS
PS
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1225 AA;
OY
OY 1 FSRYYAR 6
OY |||||
Db 613 FSRYYAR 618
Query Match 100.0%; Score 31; DB 4; Length 1225;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
RESULT 25
ABG24819
ID ABG24819 standard; protein; 1258 AA.
XX
AC ABG24819;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #24810.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS89006.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
XX
PS Claim 20; SEQ ID NO 55178; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1258 AA;
XX
Query Match 100.0%; Score 31; DB 4; Length 1258;
Best Local Similarity 100.0%; Pred. NO. 4.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FSRYAR 6
|:|:|:|
Db 1199 FSRYAR 1204
XX
RESULT 26
AAU60688
ID AAU60688 standard; protein: 58 AA.
XX
AC AAU60688;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #21584.
XX
KM SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KM uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KM inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KM dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US012865.
XX
PR 21-APR-2000; 2000US-0199047P.
PR 02-JUN-2000; 2000US-0208841P.
PR 07-JUL-2000; 2000US-0216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhactia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
XX WPI; 2001-616774/71.
DR N-PSDB; AAS59612.

XX
XX Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
XX
PS Example 1; SEQ ID NO 21883; 1069pp; English.
XX
XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 58 AA;
XX
Query Match 90.3%; Score 28; DB 4; Length 58;
Best Local Similarity 83.3%; Pred. No. 95;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 FSRYAR 6
|:|:|:|
Db 14 FARYAR 19
XX
RESULT 27
ABM57207
ID ABM57207 standard; protein: 58 AA.
XX
AC ABM57207;
XX
DT 20-OCT-2003 (first entry)
XX
DE Propionibacterium acnes predicted ORF-encoded polypeptide #21883.
XX
KM Acne vulgaris; antisephorhoetic; dermatological; antibacterial;
KM immunostimulant; immune response; vaccine.
XX
OS Propionibacterium acnes.
XX
PN WO2003033515-A1.
XX
PD 24-APR-2003.
XX
PF 11-OCT-2002; 2002WO-US032727.
XX
PR 15-OCT-2001; 2001US-00978825.
XX
PA (CORI-) CORIXA CORP.
XX
PI Mitcham JL, Skeiky YAW, Persing DH, Bhactia A, Maisonneuve JL;
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
PI Barth B, Vailieve-Douglas J;
XX
XX WPI; 2003-381789/36.
DR N-PSDB; ACF64541.
XX
XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,

PT or for stimulating an immune response specific for a *P. acnes* protein.
 XX Example 1; SEQ ID NO 21883; 1481bp; English.
 XX
 CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
 CC encoding a Propionibacterium *acnes* protein. The invention also relates to
 CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
 CC immunogenic fragments of *P. acnes* polypeptides. The invention
 CC additionally encompasses expression vectors and host cells comprising a
 CC polynucleotide of the invention; antibodies against polypeptides of the
 CC invention; fusion proteins comprising a polypeptide of the invention; a
 CC method for stimulating an immune response specific for a *P. acnes*
 CC polypeptide and an isolated T cell population comprising T cells prepared
 CC via this method; a vaccine composition (comprising *P. acnes* polypeptides,
 CC polynucleotides, antibodies, fusion proteins, T cell populations, or
 CC antigen-presenting cells that express the polypeptide); a method and kit
 CC for detecting or determining the presence or absence of *P. acnes* in a
 CC patient; and a method for inhibiting the development of *P. acnes* in a
 CC patient. The *P. acnes* polypeptides, polynucleotides, antibodies, fusion
 CC proteins, T cell populations or antigen-presenting cells that express the
 CC polypeptides are useful for diagnosing, preventing or treating *acne*
 CC vulgaris, or for stimulating an immune response specific for a *P. acnes*
 CC protein. The polynucleotides can also be used as probes or primers for
 CC nucleic acid hybridisation. The vaccine composition is useful for the
 CC stimulation of an immune response against *P. acnes*, or for treating *acne*,
 CC and the kit is useful for performing a diagnostic assay. The present
 CC sequence represents a polypeptide predicted to be encoded by an ORF (open
 CC reading frame) contained within the *P. acnes* polynucleotides of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIP0 at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences)
 CC
 XX Sequence 58 AA;
 SQ
 OY Query Match 90.3%; Score 28; DB 6; Length 58;
 Db Best Local Similarity 83.3%; Pred. No. 95; Indels 0; Gaps 0;
 Matches 5; Conservative 1; Mismatches 0;
 1 FSRYYAR 6
 14 FARYYAR 19
 RESULT 28
 AAY12203
 ID AAY12203 standard; protein; 85 AA.
 XX
 AC AAY12203;
 XX
 DT 18-JUN-1999 (first entry)
 XX
 DE Human 5' EST secreted protein SEQ ID NO: 516.
 XX
 KW Human; secreted protein; EST; expressed sequence tag; diagnosis;
 KW forensic; gene therapy; chromosome mapping; signal peptide;
 KW upstream regulatory sequence; cytokine activity; cell proliferation;
 KW differentiation; haematopoiesis regulation; tissue growth regulation;
 KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
 KW thrombolytic; anti-inflammatory; tumour inhibition.
 XX
 OS Homo sapiens.
 XX
 PN MO9906554-A2.
 XX
 PD 11-FEB-1999.
 XX
 PF 31-JUL-1998; 98WO-IB001238.
 XX
 PR 01-AUG-1997; 97US-00905134.
 XX
 PA (GSEST) GENSET.
 XX
 PI Dumas Milne Edwards J, Duclert A, Lacroix B;
 PI

XX WPI; 1999-153784/13.
 DR N-PSDB; AAX41036.
 XX
 PT New nucleic acids encoding human secreted proteins - obtained from cDNA
 PT libraries prepared from kidney, fetal kidney, dystrophic muscle, muscle
 PT and heart tissue.
 XX
 PS Claim 34; Page 585; 622pp; English.
 XX
 CC AAX40826 to AAX41093 represent 5' expressed sequence tags (ESTs) for
 CC human secreted proteins, and encode the proteins given in AAY01602 and
 CC AAY11994 to AAY12260, respectively. The proteins given represent the
 CC signal peptide and an N-terminal fragment of a secreted protein. The
 CC nucleic acid sequences can be used for producing secreted human gene
 CC products. They can also be used to develop products for diagnosis and
 CC therapy. The proteins obtained may have cytokine activity, cell
 CC proliferation/differentiation activity, haematopoiesis regulating
 CC activity, tissue growth regulating activity, reproductive hormone
 CC regulating activity, chemotactic/chemokinetic activity, haemostatic and
 CC thrombolytic activity, receptor/ligand activity, anti-inflammatory
 CC activity, tumour inhibition activity or other activities. The products
 CC can be used in forensic, gene therapy and chromosome mapping procedures.
 CC The sequences can also be used for obtaining corresponding promoter
 CC sequences. The nucleic acids encoding the signal peptide can be used for
 CC directing extracellular secretion of a polypeptide or the insertion of a
 CC polypeptide into a membrane, or importing a polypeptide into a cell
 CC
 XX Sequence 85 AA;
 SQ
 OY Query Match 90.3%; Score 28; DB 2; Length 85;
 Db Best Local Similarity 83.3%; Pred. No. 1.4e+02; Indels 0; Gaps 0;
 Matches 5; Conservative 1; Mismatches 0;
 1 FSRYYAR 6
 69 FARYYAR 74
 RESULT 29
 ABR40679
 ID ABR40679 standard; protein; 90 AA.
 XX
 AC ABR40679;
 XX
 DT 16-MAY-2003 (first entry)
 XX
 DE Oryza sativa oil trait related protein sequence SEQ ID NO:178.
 XX
 KW Plant; oil trait; oil phenotype; altered lipid profile; MAP kinase;
 KW receptor-like protein kinase; mitogen activated protein kinase; oil;
 KW lipid-like transcription factor caleosin; ATP citrate lyase; SNF1;
 KW CKC-like transcription factor; antisense inhibition; co-suppression;
 KW transgenic plant.
 XX
 OS Oryza sativa.
 XX
 PN MO2003002751-A2.
 XX
 PD 09-JAN-2003.
 XX
 PF 27-JUN-2002; 2002WO-US020152.
 XX
 PR 29-JUN-2001; 2001US-0301913P.
 XX
 PA (DHPO) DU PONT DE NEMOURS & CO E I.
 PA (PTON-) PIONEER HI-BRED INT INC.
 XX
 PI Allen SM, Allen WB, Cahoon RE, Epelbaum S, Famodu OO, Harvell LT,
 PI Jones TJ, Kinney AJ, Klein TM, Li C, Oliveira IC, Sakai H, Shen B,
 PI Tarczyński MC;
 XX
 DR WPI; 2003-201509/19.
 DR

```

DR N-PSDB; ACC00714.
XX
PT Novel nucleotide fragment encoding polypeptides having receptor-like
PT protein kinase activity, caleosin-like activity, useful for altering oil
PT phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.
XX
PS Claim 12; Page 215; 542pp; English.
XX
CC The present invention describes an isolated nucleotide fragment (I)
CC comprising a nucleic acid sequence (NS) chosen from a NS encoding a
CC polypeptide (PP) having receptor-like protein kinase activity, mitogen
CC activated protein (MAP)-kinase activity, lipid-like transcription factor
CC activity, caleosin-like activity, ATP citrate lyase activity, SNF1-like
CC activity and CKC1-like transcription factor activity. Also described: (1)
CC complement (II) of (I); (2) a chimeric construct (III) comprising (I) or
CC (II), operably linked to a regulatory sequence; (3) a plant (IV)
CC comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5)
CC oil obtained from (V). (I) or its part can be used in antisense
CC inhibition or co-suppression in a transformed plant. (III) is useful for
CC altering the oil phenotype in a plant such as corn, soybean, wheat, rice,
CC canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for
CC creating transgenic plants having altered lipid profiles. (I) can also be
CC used as a hybridisation probe. ACC00626 to ACC00868 and ABR40591 to
CC ABR40879 represent sequences used in the exemplification of the present
CC invention
CC
XX
SQ Sequence 90 AA;
XX
Query Match 90.3%; Score 28; DB 6; Length 90;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 FSRYAR 6
Db 12 FSKYAR 17
XX
RESULT 30
ABR40677
ID ABR40677 standard; protein; 132 AA.
XX
AC ABR40677;
XX
XX 16-MAY-2003 (first entry)
XX
DE Oryza sativa oil trait related protein sequence SEQ ID NO:174.
XX
KW Plant; oil trait; oil phenotype; altered lipid profile; MAP kinase;
KW receptor-like protein kinase; mitogen activated protein kinase; oil;
KW lipid-like transcription factor caleosin; ATP citrate lyase; SNF1;
KW CKC1-like transcription factor; antisense inhibition; co-suppression;
KW transgenic plant.
XX
OS Oryza sativa.
XX
XX WO2003002751-A2.
XX
XX 09-JAN-2003.
XX
XX 27-JUN-2002; 2002WO-US020152.
XX
XX 29-JUN-2001; 2001US-0301913P.
XX
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
PI Allen SM, Allen WB, Cahoon RB, Epelbaum S, Famodu OO, Harvell LT;
PI Jones TJ, Kinney AJ, Klein TM, Li C, Oliveira IC, Sakai H, Shen B;
PI Tarczyński MC;
XX
XX WPI; 2003-201509/19.
XX
XX N-PSDB; ACC00712.

```

```

PT Novel nucleotide fragment encoding polypeptides having receptor-like
PT protein kinase activity, caleosin-like activity, useful for altering oil
PT phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.
XX
PS Claim 12; Page 212; 542pp; English.
XX
CC The present invention describes an isolated nucleotide fragment (I)
CC comprising a nucleic acid sequence (NS) chosen from a NS encoding a
CC polypeptide (PP) having receptor-like protein kinase activity, mitogen
CC activated protein (MAP)-kinase activity, lipid-like transcription factor
CC activity, caleosin-like activity, ATP citrate lyase activity, SNF1-like
CC activity and CKC1-like transcription factor activity. Also described: (1)
CC complement (II) of (I); (2) a chimeric construct (III) comprising (I) or
CC (II), operably linked to a regulatory sequence; (3) a plant (IV)
CC comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5)
CC oil obtained from (V). (I) or its part can be used in antisense
CC inhibition or co-suppression in a transformed plant. (III) is useful for
CC altering the oil phenotype in a plant such as corn, soybean, wheat, rice,
CC canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for
CC creating transgenic plants having altered lipid profiles. (I) can also be
CC used as a hybridisation probe. ACC00626 to ACC00868 and ABR40591 to
CC ABR40879 represent sequences used in the exemplification of the present
CC invention
CC
XX
SQ Sequence 132 AA;
XX
Query Match 90.3%; Score 28; DB 6; Length 132;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 FSRYAR 6
Db 73 FSKYAR 78
XX
RESULT 31
AAM67901
ID AAM67901 standard; protein; 139 AA.
XX
AC AAM67901;
XX
XX 25-MAR-1999 (first entry)
XX
DE Human secreted protein encoded by gene 36 clone HODCL36.
XX
KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX FH Misc-difference 139 /label= unknown
XX FT
XX
XX WO9842738-A1.
XX
XX 01-OCT-1998.
XX
XX 19-MAR-1998; 98WO-US005311.
XX
XX 21-MAR-1997; 97US-0041276P.
XX 21-MAR-1997; 97US-0041277P.
XX 21-MAR-1997; 97US-0041281P.
XX 21-MAR-1997; 97US-0042344P.
XX 30-MAY-1997; 97US-0048069P.
XX 30-MAY-1997; 97US-0048094P.
XX 30-MAY-1997; 97US-0048095P.

```


XX MO2003048304-A2.
 XX 12-JUN-2003.
 XX 25-NOV-2002; 2002MO-US036123.
 XX 29-NOV-2001; 2001US-0333777P.
 XX 18-NOV-2002; 2002US-0426742P.
 XX (AMHP) WYETH HOLDINGS CORP.
 XX Fletcher LD, Memichael JC, Russell DP, Zagursky RJ;
 XX WPI; 2003-505284/47.
 XX N-PSDB; ADB06539.
 XX New Alloiococcus otitidis polynucleotides and polypeptides, useful for
 PT treating and diagnosing diseases, drug screening assays and monitoring of
 PT effects during drug clinical trials.
 XX Claim 33; SEQ ID NO 480; 1019p; English.
 XX The present invention describes an isolated polynucleotide (I) of
 CC Alloiococcus otitidis genomic DNA, which encodes an antigenic protein.
 CC Alloiococcus otitidis is a Gram-positive bacterium. Also described: (1)
 CC an isolated polypeptide that is encoded by the polynucleotide (I); (2) an
 CC expression vector comprising the novel isolated polynucleotide (I), its
 CC complement, degenerate variant or fragment; (3) a genetically engineered
 CC host cell, transfected, transformed or infected with the vector of (2);
 CC (4) an antibody specific for the polypeptide of (1); (5) an immunogenic
 CC composition comprising the polypeptide, its complement, biological
 CC equivalent or fragment, or the polynucleotide that is comprised in the
 CC expression vector; (6) a pharmaceutical composition comprising the
 CC polypeptide of (1) and a carrier; (7) a protein chip comprising an array
 CC of the polypeptide of (1), their biological equivalent or fragment; (8)
 CC immunising against Alloiococcus otitidis by administering to a host the
 CC immunogenic composition; (9) detecting and/or identifying Alloiococcus
 CC otitidis in the biological sample; (10) a kit comprising a container
 CC containing the novel polynucleotide, its degenerate variant or fragment,
 CC or the antibody of (4); and (11) producing a polypeptide by culturing the
 CC polypeptide from the culture. (I) can be used in gene therapy. The present
 CC polynucleotides, polypeptides, antibodies and compositions of the present
 CC invention can be used for treating and diagnosing diseases, drug
 CC screening assays and monitoring of effects during drug clinical trials.
 CC The polynucleotides are useful for expressing and detecting Alloiococcus
 CC otitidis. The present sequence represents an Alloiococcus otitidis
 CC antigen protein from the present invention.
 XX Sequence 172 AA;
 SQ
 PS Query Match 90.3%; Score 28; DB 6; Length 172;
 CC Best Local Similarity 83.3%; Pred. No. 2.9e+02;
 PT Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 XX
 QY 1 FSRYSR 6
 DB 149 FSRYSR 154
 XX
 XX RESULT 34
 XX AAW67842
 XX ID AAW67842 standard; protein; 174 AA.
 XX
 XX AAW67842;
 XX
 XX 25-MAR-1999 (first entry)
 XX
 XX Human secreted protein encoded by gene 36 clone HODCU36.
 XX
 XX Human; secreted protein; fusion protein; gene therapy; protein therapy;
 KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;

KW developmental abnormality; foetal deficiency; blood; allergy; renal;
 KW immune system; aschma; lymphocytic disease; brain; hepatic; lymphoma;
 KW inflammation; ischemic shock; Alzheimer's disease; resectosis; AIDS;
 KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
 KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid digestion;
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
 XX
 OS Homo sapiens.
 XX
 XX W09842738-A1.
 XX
 XX 01-OCT-1998.
 XX
 XX 19-MAR-1998; 98MO-US005311.
 XX
 XX 21-MAR-1997; 97US-0041276P.
 XX 21-MAR-1997; 97US-0041277P.
 XX 21-MAR-1997; 97US-0041281P.
 XX 21-MAR-1997; 97US-0042344P.
 XX 30-MAY-1997; 97US-0048069P.
 XX 30-MAY-1997; 97US-0048094P.
 XX 30-MAY-1997; 97US-0048095P.
 XX 30-MAY-1997; 97US-0048096P.
 XX 30-MAY-1997; 97US-0048131P.
 XX 30-MAY-1997; 97US-0048135P.
 XX 30-MAY-1997; 97US-0048154P.
 XX 30-MAY-1997; 97US-0048160P.
 XX 30-MAY-1997; 97US-0048186P.
 XX 30-MAY-1997; 97US-0048187P.
 XX 30-MAY-1997; 97US-0048188P.
 XX 30-MAY-1997; 97US-0048351P.
 XX 30-MAY-1997; 97US-0048352P.
 XX 30-MAY-1997; 97US-0048355P.
 XX 30-MAY-1997; 97US-0050937P.
 XX 05-AUG-1997; 97US-0054804P.
 XX 19-AUG-1997; 97US-0056370P.
 XX 02-OCT-1997; 97US-0060862P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Young P, Greene JM, Ferrie AM, Ruben SM, Rosen CA, Duan R, Hu J;
 PI Florence KA, Olsen HS, Ebner R, Brewer LA, Moore PA, Shi Y;
 PI Lafleur DM, Ni J;
 XX WPI; 1999-070066/06.
 DR N-PSDB; AAX00646.
 XX
 XX New isolated human genes and the secreted polypeptides they encode -
 PT useful for diagnosis and treatment of e.g. cancers, neurological
 PT disorders, immune diseases, inflammation or blood disorders.
 XX
 PS Claim 11; Page 291; 385p; English.
 XX
 XX This sequence represents a secreted human protein encoded by the gene
 CC clone detailed in the descriptor line. The gene can be used to generate
 CC fusion proteins by linking to the gene to a human immunoglobulin Fc
 CC portion (e.g. AAX00602) for increasing the stability of the fused protein
 CC as compared to the human protein only. The invention relates to 87 novel
 CC genes and their fragments (nucleic acid sequences: AAX00611-X00724; amino
 CC acid sequences AAW67807-W68004) which are useful for preventing, treating
 CC or ameliorating medical conditions e.g. by protein or gene therapy. Also,
 CC pathological conditions can be diagnosed by determining the amount of the
 CC new polypeptides in a sample or by determining the presence of mutations
 CC in the new polynucleotides. Specific uses are described for each of the
 CC 87 polynucleotides, based on which tissues they are most highly expressed
 CC in (see AAX00611 for described uses)
 XX
 XX Sequence 174 AA;
 SQ
 PS Query Match 90.3%; Score 28; DB 2; Length 174;
 CC Best Local Similarity 83.3%; Pred. No. 2.9e+02;
 PT Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 XX

OY 1 FSRYPAR 6
|:||||
XX
DB 69 FARYPAR 74

RESULT 35

ADY14896
ID ADY14896 standard; protein; 174 AA.

AC ADY14896;

DT 05-MAY-2005 (first entry)

DE PRO polypeptide SEQ ID NO 702.

XX Antiflammatory; Immune disorder; Dermatological; Immunosuppressive;
XX Antineumatic; Antiarthritic; Osteopathic; Hemostatic; Antineumatic;
XX Antitumor; Antidiabetic; Nephrotropic; CNS-Gen.; Hepatotropic;
XX Vitucide; Gastrointestinal-Gen.; Antipsoriatic; Antialsthetic;
XX Antiallergic; ds; gene; diagnosis.

XX Homo sapiens.

PN WO2005016962-A2.

PD 24-FEB-2005.

PF 11-AUG-2004; 2004WO-US026249.

PR 11-AUG-2003; 2003US-0493546P.

PA (GETH) GENENTECH INC.

PI Abbas A, Clark H, Ouyang W, Williams MP, Wood WI, Wu TD;

DR WPI; 2005-182330/19.

XX New nucleic acid encoding PRO polypeptide, useful for diagnosing and
XX creating an immune related disorder, e.g. systemic lupus erythematosus,
XX rheumatoid arthritis, osteoarthritis, thyroiditis, or diabetes mellitus.
XX Claim 8; SEQ ID NO 702; 158pp; English.

XX The invention relates to an isolated nucleic acid encoding a PRO
XX polypeptide. The polypeptide, agonist or an antagonist, antibody,
XX composition, and method are useful for diagnosing and treating an immune
XX related disorder, e.g. systemic lupus erythematosus, rheumatoid
XX arthritis. The present sequence represents a DNA encoding a PRO
XX polypeptide.

XX Sequence 174 AA;

Query Match 90.3%; Score 28; DB 9; Length 174;

Best Local Similarity 83.3%; Pred. No. 2.9e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FSRYPAR 6
|:||||
DB 69 FARYPAR 74

RESULT 36
AEF69875
ID AEF69875 standard; protein; 174 AA.

XX AEF69875;

DT 06-APR-2006 (first entry)

DE Microsatellite instability colorectal tumor marker protein SEQ ID NO:53.
XX satellite DNA, tumor marker; therapeutic; diagnosis; prognosis;
KW

KW colorectal tumor; cancer; neoplasm; cytostatic.

XX Homo sapiens.

XX WO2006015047-A2.

XX 09-FEB-2006.

XX 27-JUL-2005; 2005WO-US026647.

XX 28-JUL-2004; 2004US-0591969P.

XX 13-APR-2005; 2005US-00105233.

XX (FARB) BAYER HEALTHCARE LLC.

PA (MAYO-) MAYO FOUND MEDICAL EDUCATION & RES.

PI Bugess C, Catino TJ, Molino G, Myerow S, Thiagalingam A;

DR WPI; 2006-155744/16.

DR N-PSDB; AEF69841.

PT Detecting microsatellite instability, useful for monitoring the onset,
PT progression, or regression of cancer, comprises measuring the expression
PT of genes.

PS Claim 24; SEQ ID NO 53; 282pp; English.

XX The invention relates to a method for monitoring the onset, progression,
XX or regression of cancer or a pre-malignant condition, particularly
XX colorectal cancer, in an individual comprising detecting microsatellite
XX instability in an individual by measuring the expression of the sequence
XX of one or more nucleotide sequences defined in the specification, where
XX microsatellite instability is determined as overexpression of the
XX sequence by at least 2-fold compared to the expression in an individual
XX known to not have microsatellite instability. Alternatively, the method
XX comprises detecting binding of a ligand to a polypeptide encoded by one
XX or more of the above nucleotide sequences. The methods are useful for
XX detecting microsatellite instability in an individual, for identifying
XX colon tissue having microsatellite instability, for monitoring the onset,
XX progression, or regression of cancer or a pre-malignant condition in a
XX subject, for determining prognosis for cancer or a pre-malignant
XX condition in a subject. The nucleotides and polypeptides are useful for
XX determining the efficacy of a test compound or a therapy, for inhibiting
XX cancer in a subject and for selecting a composition for inhibiting cancer
XX in a subject. The compound, composition and method are useful for
XX treating cancer. This sequence is a protein encoded by a marker DNA
XX sequence specific for a high degree of microsatellite instability (MSI-H
XX marker) relative to colorectal tumors.

XX Sequence 174 AA;

Query Match 90.3%; Score 28; DB 10; Length 174;

Best Local Similarity 83.3%; Pred. No. 2.9e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FSRYPAR 6
|:||||
DB 69 FARYPAR 74

RESULT 37
AEA17517
ID AEA17517 standard; protein; 243 AA.

XX AEA17517;

DT 28-JUL-2005 (first entry)

DE Thale cress caleosin-related family protein, SEQ ID NO: 122.
XX Apolipoprotein; expression; transgenic plant; pharmaceutical;
XX vascular disease; atherosclerosis; vasotropic; antiarteriosclerotic;
KW

KM caleosin.
 XX Arabidopsis thaliana.
 XX
 PN WO2005047455-A2.
 XX
 PD 26-MAY-2005.
 XX
 PF 15-NOV-2004; 2004WO-CA001960.
 XX
 PR 14-NOV-2003; 2003US-0519606P.
 PR 16-JUN-2004; 2004US-0579733P.
 XX
 PA (SEMB-) SEMBIOSYS GENETICS INC.
 XX
 PI Moloney MM, Reid A;
 XX
 DR WPI; 2005-372355/38.
 XX
 PT Expressing apolipoprotein in plants which may be used for treating
 PT vascular diseases, comprises introducing a chimeric nucleic acid
 PT construct into a plant cell.
 XX
 PS Disclosure; SEQ ID NO 122; 530pp; English.
 XX
 CC The present invention relates to methods for expressing apolipoprotein in
 CC plants by introducing a chimeric nucleic acid construct into a plant
 CC cell. Specifically, this comprises providing a chimeric nucleic acid
 CC construct comprising a nucleic acid sequence capable of controlling
 CC expression in plant cells; and a sequence encoding an apolipoprotein
 CC polypeptide; introducing the chimeric nucleic acid construct into a plant
 CC cell; and growing the plant cell into a mature plant where the plant
 CC expresses apolipoprotein. Also disclosed is a composition of
 CC substantially pure apolipoprotein obtained from a plant, or substantially
 CC pure oil bodies comprising an apolipoprotein obtained from plants. The
 CC nucleic acid sequence capable of controlling expression is a seed-
 CC preferred or constitutive promoter such as a phaseolin promoter and a
 CC ubiquitin promoter. The chimeric nucleic acid construct additionally
 CC comprises a nucleic acid sequence encoding a stabilizing polypeptide
 CC linked in frame to the sequence encoding apolipoprotein. It is a plant-
 CC specific oil body protein or a thioredoxin (SEQ ID NO: 56). The oil body
 CC protein is selected from oleosin, caleosin, and steroleosin. The non-
 CC plant specific stabilizing polypeptide is a green fluorescent protein
 CC (SEQ ID NO: 57) or a single chain antibody. The single-chain antibody is
 CC capable of facilitating purification of apolipoprotein expressed in seeds
 CC or is capable of specifically associating with an oil body protein.
 CC Alternatively, the single-chain antibody is a single-chain Fv antibody
 CC (SEQ ID NO: 240) capable of associating with the 18 kDa oleosin from *A.*
 CC thaliana (D9scrV). The pharmaceutical apolipoprotein formulation may be
 CC used to treat vascular diseases, such as atherosclerosis. The present
 CC sequence is a thale cress caleosin protein sequence.
 XX
 SQ Sequence 243 AA;
 XX
 QY Query Match 90.3%; Score 28; DB 9; Length 243;
 Best Local Similarity 83.3%; Pred. No. 4.1e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 XX
 DB 1 FSRVAR 6
 159 FSKYAR 164
 XX
 RESULT 38
 ABR40670
 ID ABR40670 standard; protein; 244 AA.
 XX
 AC ABR40670;
 XX
 DT 16-MAY-2003 (first entry)
 XX
 DE Zea mays oil trait related protein sequence SEQ ID NO:160.
 XX

KM Plant; oil trait; oil phenotype; altered lipid profile; MAP kinase;
 KM receptor-like protein kinase; mitogen activated protein kinase; oil;
 KM lipo-5-like transcription factor caleosin; ATP citrate lyase; SNF1;
 KM CKC-like transcription factor; antisense inhibition; co-suppression;
 KM transgenic plant.
 XX
 OS Zea mays.
 XX
 PN WO2003002751-A2.
 XX
 PD 09-JAN-2003.
 XX
 PF 27-JUN-2002; 2002WO-US020152.
 XX
 PR 29-JUN-2001; 2001US-0301913P.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E I.
 PA (PION-) PIONEER HI-BRED INT INC.
 XX
 PI Allen SM, Allen WB, Caboon RE, Epelbaum S, Famodu OO, Harvey LL,
 PI Jones TJ, Kinney AJ, Klein TM, Li C, Oliveira IC, Sakai H, Shen B,
 PI Tarczyński MC;
 XX
 DR WPI; 2003-201509/19.
 DR N-PSDB; ACC00705.
 XX
 PT Novel nucleotide fragment encoding polypeptides having receptor-like
 PT protein kinase activity, caleosin-like activity, useful for altering oil
 PT phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.
 XX
 PS Claim 12; Page 202-203; 542pp; English.
 XX
 CC The present invention describes an isolated nucleotide fragment (I)
 CC comprising a nucleic acid sequence (NS) chosen from a NS encoding a
 CC polypeptide (PP) having receptor-like protein kinase activity, mitogen
 CC activated protein (MAP)-kinase activity, lipo-5-like transcription factor
 CC activity, caleosin-like activity, ATP citrate lyase activity, SNF1-like
 CC activity and CKC-like transcription factor activity. Also described: (1)
 CC complement (II) of (I); (2) a chimeric construct (III) comprising (I) or
 CC (II), operably linked to a regulatory sequence; (3) a plant (IV)
 CC comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5)
 CC oil obtained from (V). (I) or its part can be used in antisense
 CC inhibition or co-suppression in a transformed plant. (III) is useful for
 CC altering the oil phenotype in a plant such as corn, soybean, wheat, rice,
 CC canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for
 CC creating transgenic plants having altered lipid profiles. (I) can also be
 CC used as a hybridisation probe. ACC00626 to ACC00868 and ABR40591 to
 CC ABR40879 represent sequences used in the exemplification of the present
 CC invention.
 XX
 SQ Sequence 244 AA;
 XX
 QY Query Match 90.3%; Score 28; DB 6; Length 244;
 Best Local Similarity 83.3%; Pred. No. 4.2e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 XX
 DB 1 FSRVAR 6
 163 FSKYAR 168
 XX
 RESULT 39
 ADX91903
 ID ADX91903 standard; protein; 256 AA.
 XX
 AC ADX91903;
 XX
 DT 21-APR-2005 (first entry)
 XX
 DE Plant full length insert polypeptide seqid 54567.
 XX
 DE plant protectant; plant growth regulant; gene therapy; plant;
 KM recombinant DNA construct; physical array; plant breeding marker;
 KM

KM cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
 KM extreme osmotic condition; pathogen tolerance; pest tolerance;
 KM growth rate; cell cycle pathway; disease resistance;
 KM galactomannan production; lignin production; plant growth regulator;
 KM yield; plant growth; plant development; seed oil; protein yield;
 KM protein content.
 XX
 OS unidentified.
 XX
 PN US2004034888-A1.
 XX
 PD 19-FEB-2004.
 XX
 PF 28-APR-2003; 2003US-00425114.
 XX
 PR 06-MAY-1999; 99US-00304517.
 PR 05-NOV-2001; 2001US-00985678.
 XX
 PA (LIU/) LIU J.
 PA (ZHOU/) ZHOU Y.
 PA (KOVA/) KOVALIC D K.
 PA (SCRE/) SCREEN S E.
 PA (TABA/) TABASKA J E.
 PA (CAO/) CAO Y.
 XX
 PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
 XX
 DR WPI; 2004-180133/17.
 XX
 PT New recombinant DNA construct, useful for improving plant tolerance to
 PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
 PT pests, for conferring increased resistance to plant disease, or for
 PT improving yield.
 XX
 PS Claim 1; SEQ ID NO 54567; 15pp; English.
 XX
 CC The invention describes a recombinant DNA construct comprising a
 CC polynucleotide consisting of a sequence encoding an amino acid sequence
 CC available in electronic form from the US patent office at
 CC ftp.secdat.uspto.gov/sequence.html?docID:2004034888. The polynucleotide
 CC of the invention are also useful in physical arrays of molecules and as
 CC plant breeding markers. The recombinant DNA construct is useful for
 CC improving plant tolerance to cold, heat, drought, herbicides, extreme
 CC osmotic conditions, pathogens or pests, for manipulating growth rate in
 CC plant cells by modification of the cell cycle pathway, for conferring
 CC increased resistance to plant disease, for producing galactomannan,
 CC lignin or plant growth regulators, for increasing the rate of homologous
 CC recombination in plants, for improving yield by modification of
 CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
 CC or by providing improved plant growth and development under at least one
 CC stress condition or for modifying seed oil or protein yield and/or
 CC content. This is the amino acid sequence of a plant full length insert
 CC polypeptide that can be used in the recombinant DNA construct of the
 CC invention.
 XX
 SQ Sequence 256 AA;
 XX
 QY Query Match 90.3%; Score 28; DB 8; Length 256;
 Best Local Similarity 83.3%; Pred. No. 4.4e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 XX
 DB 1 FSRVAR 6
 174 FSKVAR 179
 XX
 RESULT 40
 ABG07163
 ID ABG07163 standard; protein; 301 AA.
 XX
 AC ABG07163;
 XX
 DT 13-FEB-2002 (first entry)

XX
 DE Novel human diagnostic protein #7154.
 XX
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KM food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US008631.
 XX
 PR 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR N-PSDB; AAS71350.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 20; SEQ ID NO 37522; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 301 AA;
 XX
 QY Query Match 90.3%; Score 28; DB 4; Length 301;
 Best Local Similarity 83.3%; Pred. No. 5.2e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 XX
 DB 1 FSRVAR 6
 90 FARVAR 95
 XX
 RESULT 41
 AEA17523
 ID AEA17523 standard; protein; 301 AA.
 XX
 AC AEA17523;
 XX
 DT 28-JUL-2005 (first entry)
 XX
 DE Thale cress caleosin-1 protein, SEQ ID NO: 128.

XX	Apolipoprotein; expression; transgenic plant; pharmaceutical;
KM	vascular disease; atherosclerosis; vasotropic; antiarteriosclerotic;
XX	caleosin.
XX	
OS	Arabidopsis thaliana.
XX	
PN	WO2005047455-A2.
PD	
XX	26-MAY-2005.
PF	
XX	15-NOV-2004; 2004WO-CA001960.
PR	
XX	14-NOV-2003; 2003US-0519606P.
PR	16-JUN-2004; 2004US-0579733P.
XX	
PA	(SEMB-) SEMBIOSYS GENETICS INC.
XX	
P1	Moloney MM, Reid A;
XX	
DR	WPI; 2005-372355/38.
XX	
PT	Expressing apolipoprotein in plants which may be used for treating
PT	vascular diseases, comprises introducing a chimeric nucleic acid
XX	construct into a plant cell.
PS	
XX	Disclosure; SEQ ID NO 128; 530pp; English.
XX	
CC	The present invention relates to methods for expressing apolipoprotein in
CC	plants by introducing a chimeric nucleic acid construct into a plant
CC	cell. Specifically, this comprises providing a chimeric nucleic acid
CC	construct comprising a nucleic acid sequence capable of controlling
CC	expression in plant cells; and a sequence encoding an apolipoprotein
CC	polypeptide; introducing the chimeric nucleic acid construct into a plant
CC	cell; and growing the plant cell into a mature plant where the plant
CC	expresses apolipoprotein. Also disclosed is a composition of
CC	substantially pure apolipoprotein obtained from a plant, or substantially
CC	pure oil bodies comprising an apolipoprotein obtained from plants. The
CC	nucleic acid sequence capable of controlling expression is a seed-
CC	preferred or constitutive promoter such as a phaseolin promoter and a
CC	ubiquitin promoter. The chimeric nucleic acid construct additionally
CC	comprises a nucleic acid sequence encoding a stabilizing polypeptide
CC	linked in frame to the sequence encoding apolipoprotein. It is a plant-
CC	specific oil body protein or a thioredoxin (SEQ ID NO: 56). The oil body
CC	protein is selected from oleosin, caleosin, and steroleosin. The non-
CC	plant specific stabilizing polypeptide is a green fluorescent protein
CC	(SEQ ID NO: 57) or a single chain antibody. The single-chain antibody is
CC	capable of facilitating purification of apolipoprotein expressed in seeds
CC	or is capable of specifically associating with an oil body protein.
CC	Alternatively, the single-chain antibody is a single-chain Fv antibody
CC	(SEQ ID NO: 240) capable of associating with the 18 kDa oleosin from A.
CC	thaliana (D9scFv). The pharmaceutical apolipoprotein formulation may be
CC	used to treat vascular diseases, such as atherosclerosis. The present
CC	sequence is a thale cress caleosin protein sequence.
XX	
SQ	Sequence 301 AA;
XX	
QY	Query Match 90.3%; Score 28; DB 9; Length 301;
DB	Best Local Similarity 83.3%; Pred. No. 5.2e+02;
XX	Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX	
XX	1 FSRVAR 6
XX	
XX	220 FSKYAR 225
XX	
XX	RESULT 42
ID	AEAI7522
XX	AEAI7522 standard; protein; 301 AA.
XX	
XX	AEAI7522;
DT	28-JUL-2005 (first entry)

```

DE   Thale cress caleosin-1 protein, SEQ ID NO: 127.
XX
XX   Apolipoprotein; expression; transgenic plant; pharmaceutical;
XX   vascular disease; atherosclerosis; vasotropic; antiarteriosclerotic;
XX   caleosin.
XX   Arabidopsis thaliana.
XX   OS
XX   PN
XX   WO2005047455-A2.
XX   PD
XX   26-MAY-2005.
XX   PF
XX   15-NOV-2004; 2004WO-CA001960.
XX   PR
XX   14-NOV-2003; 2003US-0519606P.
XX   PR   16-JUN-2004; 2004US-0579733P.
XX   PA
XX   (SEMB-) SEMBIOSYS GENETICS INC.
XX   PI
XX   Moloney MM, Reid A;
XX   WPI; 2005-372355/38.
XX   DD
XX   Expressing apolipoprotein in plants which may be used for treating
XX   PT   vascular diseases, comprises introducing a chimeric nucleic acid
XX   PT   construct into a plant cell.
XX
XX   Disclosure; SEQ ID NO 127; 530pp; English.
XX
XX   The present invention relates to methods for expressing apolipoprotein in
XX   CC   plants by introducing a chimeric nucleic acid construct into a plant
XX   CC   cell. Specifically, this comprises providing a chimeric nucleic acid
XX   CC   construct comprising a nucleic acid sequence capable of controlling
XX   CC   expression in plant cells, and a sequence encoding an apolipoprotein
XX   CC   polypeptide; introducing the chimeric nucleic acid construct into a plant
XX   CC   cell; and growing the plant cell into a mature plant where the plant
XX   CC   expresses apolipoprotein. Also disclosed is a composition of
XX   CC   substantially pure apolipoprotein obtained from a plant, or substantially
XX   CC   pure oil bodies comprising an apolipoprotein obtained from plants. The
XX   CC   nucleic acid sequence capable of controlling expression is a seed-
XX   CC   preferred or constitutive promoter such as a phaseolin promoter and a
XX   CC   ubiquitin promoter. The chimeric nucleic acid construct additionally
XX   CC   comprises a nucleic acid sequence encoding a stabilizing polypeptide
XX   CC   linked in frame to the sequence encoding apolipoprotein. It is a plant-
XX   CC   specific oil body protein or a thiorodoxin (SEQ ID NO: 56). The oil body
XX   CC   protein is selected from oleosin, caleosin, and steroleosin. The non-
XX   CC   plant specific stabilizing polypeptide is a green fluorescent protein
XX   CC   (SEQ ID NO: 57) or a single chain antibody. The single-chain antibody is
XX   CC   capable of facilitating purification of apolipoprotein expressed in seeds
XX   CC   or is capable of specifically associating with an oil body protein.
XX   CC   Alternatively, the single-chain antibody is a single-chain Fv antibody
XX   CC   (SEQ ID NO: 240) capable of associating with the 18 kDa oleosin from A.
XX   CC   thaliana (D9serfV). The pharmaceutical apolipoprotein formulation may be
XX   CC   used to treat vascular diseases, such as atherosclerosis. The present
XX   CC   sequence is a thale cress caleosin protein sequence.
XX   SQ
XX
XX   Sequence 301 AA;
XX
XX   Query Match: 90.3%; Score 28; DB 9; Length 301;
XX   Beet Local Similarity 83.3%; Pred. NO. 5.2e+02;
XX   Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0.
XX
XX   OY 1 FSRVAR 6
XX   ||:|||
XX   Db 220 FSKVAR 225
XX
XX   RESULT 43
XX   ABG27724
XX   ID ABG27724 standard; protein; 499 AA.
XX
XX   ABG27724;

```

XX 18-FEB-2002 (first entry)
 XX
 DT
 DE Novel human diagnostic protein #27715.
 XX
 DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KM food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US008631.
 XX
 PR 31-MAR-2000; 2000US-00540217.
 XX
 PR 23-AUG-2000; 2000US-00649167.
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR N-PSDB; AAS91911.
 XX
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 20; SEQ ID NO 58083; 103bp; English.
 XX
 XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 499 AA;
 SQ
 Query Match 90.3%; Score 28; DB 4; Length 499;
 Best Local Similarity 83.3%; Pred. No. 8.7e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FSRYAR 6
 | : | | | |
 Db 89 FARVAR 94

RESULT 44
 ABG29681
 ID ABG29681 standard; protein; 520 AA.
 XX
 AC ABG29681;
 XX
 DT 18-FEB-2002 (first entry)

XX Novel human diagnostic protein #29672.
 DE
 XX
 KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KM food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US008631.
 XX
 PR 31-MAR-2000; 2000US-00540217.
 XX
 PR 23-AUG-2000; 2000US-00649167.
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR N-PSDB; AAS93868.
 XX
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 20; SEQ ID NO 60040; 103bp; English.
 XX
 XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 520 AA;
 SQ
 Query Match 90.3%; Score 28; DB 4; Length 520;
 Best Local Similarity 83.3%; Pred. No. 9e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FSRYAR 6
 | : | | | |
 Db 46 FNRVAR 51

RESULT 45
 ADN18185
 ID ADN18185 standard; protein; 681 AA.
 XX
 AC ADN18185;
 XX
 DT 02-DEC-2004 (first entry)
 DE Bacterial polypeptide #838.

XX Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
OS Bacteria.
XX US200323675-A1.
XX 18-DEC-2003.
XX 20-FEB-2003; 2003US-00369493.
XX 21-FEB-2002; 2002US-0360039P.
XX (CAOY/) CAO Y.
XX (HINK/) HINKLE G J.
XX (SLAT/) SLATER S C.
XX (CHEN/) CHEN X.
XX (GOLD/) GOLDMAN B S.
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
PI WPI; 2004-061375/06.
XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX Claim 1; SEQ ID NO 838; 122pp; English.
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX SQ Sequence 681 AA;
Query Match 90.3%; Score 28; DB 8; Length 681;
Best Local Similarity 83.3%; Pred. No. 1.2e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 FSRYAR 6
DB 515 FARYAR 520

ID ABG04036 standard; protein; 847 AA.
XX AC ABG04036;
XX DT 13-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #4027.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US0008631.
XX 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
PI WPI; 2001-639362/73.
XX N-PSDB; AAS68223.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX Claim 20; SEQ ID NO 34395; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping.
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activity. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
XX amino acid sequences of the invention. Note: The sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_ptc_sequences
SQ Sequence 847 AA;
Query Match 90.3%; Score 28; DB 4; Length 847;
Best Local Similarity 83.3%; Pred. No. 1.5e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 FSRYAR 6
DB 835 FARYAR 840

AC ABP76771;
XX
XX 07-MAR-2003 (first entry)
XX
DE N. gonorrhoeae amino acid sequence SEQ ID 72.
XX
XX Antibacterial; infection; vaccine; gene therapy.
XX
OS Neisseria gonorrhoeae.
XX
PN WO200279243-A2.
XX
PD 10-OCT-2002.
XX
XX 12-FEB-2002; 2002WO-IB002069.
XX
XX 12-FEB-2001; 2001GB-00003424.
XX
PA (CHIR-) CHIRON SPA.
XX
PI Fontana MR, Pizza M, Maignani V, Monaci E;
XX
DR WPI: 2003-058415/05.
XX
DR N-PSDB; ABZ37741.
XX
PT New protein from Neisseria gonorrhoeae, useful for the manufacture of a
XX
PT medicament for treating or preventing N. gonorrhoeae infection.
XX
PS Claim 1; Page 182; 815pp; English.
XX
XX The present invention relates to proteins from Neisseria gonorrhoeae.
CC Also disclosed are the nucleic acid molecules encoding the proteins and
CC antibodies that specifically bind to the proteins. The composition
CC comprising the protein, nucleic acid or antibody is useful for the
CC manufacture of a medicament for treating or preventing N. gonorrhoeae
CC infection, this may be in the form of a vaccine or gene therapy.
CC Sequences given in records ABP76736-ABP81046 represent nucleic acid
CC molecules of the invention
XX
SQ Sequence 1206 AA;

Query Match 90.3%; Score 28; DB 6; Length 1206;
Best Local Similarity 83.3%; Pred. No. 2.1e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FSRYYAR 6
DB 939 FARYYAR 944

RESULT 48
ABP78681
ID ABP78681 standard; protein; 1206 AA.
XX
AC ABP78681;
XX
XX 07-MAR-2003 (first entry)
XX
XX N. gonorrhoeae amino acid sequence SEQ ID 3892.
DE
XX Antibacterial; infection; vaccine; gene therapy.
XX
XX Neisseria gonorrhoeae.
XX
XX WO200279243-A2.
XX
XX 10-OCT-2002.
XX
XX 12-FEB-2002; 2002WO-IB002069.
XX
XX 12-FEB-2001; 2001GB-00003424.
XX
PA (CHIR-) CHIRON SPA.

XX
XX Fontana MR, Pizza M, Maignani V, Monaci E;
PI
XX
XX WPI: 2003-058415/05.
DR
XX N-PSDB; ABZ39651.
XX
XX New protein from Neisseria gonorrhoeae, useful for the manufacture of a
PT medicament for treating or preventing N. gonorrhoeae infection.
XX
XX Disclosure; Page 470; 815pp; English.
XX
XX The present invention relates to proteins from Neisseria gonorrhoeae.
CC Also disclosed are the nucleic acid molecules encoding the proteins and
CC antibodies that specifically bind to the proteins. The composition
CC comprising the protein, nucleic acid or antibody is useful for the
CC manufacture of a medicament for treating or preventing N. gonorrhoeae
CC infection, this may be in the form of a vaccine or gene therapy.
CC Sequences given in records ABP76736-ABP81046 represent nucleic acid
CC molecules of the invention
XX
SQ Sequence 1206 AA;

Query Match 90.3%; Score 28; DB 6; Length 1206;
Best Local Similarity 83.3%; Pred. No. 2.1e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FSRYYAR 6
DB 939 FARYYAR 944

RESULT 49
ABO69413
ID ABO69413 standard; protein; 1778 AA.
XX
XX ABO69413;
XX
XX 29-JUL-2004 (first entry)
XX
XX Pseudomonas aeruginosa polypeptide #1588.
DE
XX Pseudomonas aeruginosa infection; antibacterial.
XX
XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
XX Pseudomonas aeruginosa.
OS
XX US6551795-B1.
XX
XX 22-APR-2003.
XX
XX 18-FEB-1999; 99US-00252991.
XX
XX 18-FEB-1998; 98US-0074788P.
XX
XX 27-JUL-1998; 98US-0094130P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
PI
XX WPI: 2003-615309/58.
XX
XX N-PSDB; ABD02984.
XX
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
XX Disclosure; SEQ ID NO 18159; 455pp; English.
XX
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide
CC for the ability to bind a P. aeruginosa nucleic acid, as components of

CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences AB067826-
CC AB084396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
SQ Sequence 1778 AA;

Query Match 90.3%; Score 28; DB 7; Length 1778;

Best Local Similarity 83.3%; Pred. No. 3.2e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRYAR 6
|||:|
Db 1131 FSKYAR 1136

RESULT 50

AB878221 ID ABB78221 standard; protein; 2151 AA.

AC ABB78221;

DT 25-NOV-2002 (first entry)

DE AlphaIC subunit of a voltage-dependent L-type calcium channel.

KM Zebrafish; alphaIC subunit; voltage-dependent L-type calcium channel;
KW heart disease; cardiac arrhythmia; atrial fibrillation.

XX Danilo rerio.

OS WO200262204-A2.

PN 15-AUG-2002.

XX 06-FEB-2002; 2002WO-US003441.

PF 07-FEB-2001; 2001US-0267332P.

XX (GEHO) GEN HOSPITAL CORP.

PA Fishman MC, Rottbauer W;

PI WPI, 2002-643382/69.

XX N-PSDB; ABQ78616.

XX Determining whether a subject has or is at risk of developing disease
PT related to alphaIC subunit of voltage-dependent L-type calcium channel,
PT by detecting mutations in genes encoding the subunit in sample of
PT subject.

PS Disclosure; Page 57-61; 61pp; English.

XX The present sequence represents a zebrafish alphaIC subunit of a voltage-
CC dependent L-type calcium channel. The specification describes a method
CC for determining whether a test subject has or is at risk of developing a
CC disease or condition related to the alphaIC subunit. The method involves
CC analyzing a nucleic acid molecule sample from the test subject to
CC determine whether the test subject has a mutation in a gene encoding the
CC alphaIC subunit, where the presence of a mutation indicates that subject
CC has or is at risk of developing the disease. The method is useful for
CC determining whether a test subject (preferably mammal e.g. human) has, or
CC is at risk of developing, a disease or condition related to the alphaIC,
CC e.g. heart disease including cardiac arrhythmia or atrial fibrillation
SQ Sequence 2151 AA;

Query Match 90.3%; Score 28; DB 5; Length 2151;
Best Local Similarity 83.3%; Pred. No. 3.9e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRYAR 6
|||:|
Db 515 FSRYSR 520

Search completed: August 29, 2006, 05:59:43
Job time : 55.2079 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2006, 06:00:58 ; Search time 6 Seconds
(without alignments)
96.217 Million cell updates/sec

Title: US-10-541-343-2
Perfect score: 31
Sequence: 1 FSRVAR 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 500 summaries

Database : PIR 80: *
1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	31	100.0	356	2	C64890	probable electron
2	31	100.0	357	2	C72022	UDP-N-acetylglucos
3	31	100.0	357	2	F86603	peptidoglycan tran
4	31	100.0	463	2	C69997	probable proline t
5	31	100.0	474	2	H65051	6-phospho-beta-glu
6	31	100.0	474	2	D91075	6-phospho-beta-glu
7	31	100.0	474	2	C85920	6-phospho-beta-glu
8	31	100.0	496	2	B83591	probable transport
9	28	90.3	239	2	T07092	Ca+2-binding EF ha
10	28	90.3	304	2	T16828	hypothetical protei
11	28	90.3	409	2	A93271	probable allantoin
12	28	90.3	437	2	H95857	probable deaminase
13	28	90.3	475	2	A95376	probable ArcD1 arg
14	28	90.3	573	2	D90202	methionyl-tRNA syn
15	28	90.3	681	2	E64908	peptidyl-dipeptida
16	28	90.3	681	2	C90897	dipeptidyl carboxy
17	28	90.3	681	2	D85720	dipeptidyl carboxy
18	28	90.3	877	2	AC2211	heme transport pro
19	28	90.3	1746	2	D83181	hypothetical protei
20	27	87.1	36	2	F84791	hypothetical protei
21	27	87.1	43	2	A10982	hypothetical protei
22	27	87.1	108	2	F97521	VC03 protein homo
23	27	87.1	108	2	AH2740	conserved hypothet
24	27	87.1	208	2	B84920	hypothetical protei
25	27	87.1	220	2	B49547	nucleoside-diphosp
26	27	87.1	232	2	UC4755	ribosomal protein
27	27	87.1	286	2	AE3245	hypothetical protei
28	27	87.1	286	2	T03434	probable transport
29	27	87.1	292	2	S35983	proline transport

30	27	87.1	293	2	E87697	dUDP-4-dehydroxam
31	27	87.1	302	2	T45890	splicing factor-1i
32	27	87.1	341	2	JC5855	polyketide synthas
33	27	87.1	354	2	C83575	fructose-1,6-bisph
34	27	87.1	410	2	E75080	3-phosphoglycerate
35	27	87.1	410	2	D71065	probable phosphogl
36	27	87.1	411	2	T15705	hypothetical prote
37	27	87.1	453	2	D87672	glutamate-cysteine
38	27	87.1	456	2	AH0551	proline-specific p
39	27	87.1	457	2	B64769	proline transport
40	27	87.1	457	2	D90685	proline permease t
41	27	87.1	457	2	H85535	proline permease t
42	27	87.1	463	2	A10388	probable proline-s
43	27	87.1	465	2	C42603	phospho-beta-gluco
44	27	87.1	478	2	C87351	aldehyde dehydroge
45	27	87.1	579	2	B86565	oligopeptide perme
46	27	87.1	579	2	C72059	peptide ABC transp
47	27	87.1	605	2	S16561	glutamine-fructose
48	27	87.1	608	1	S01040	glutamine-fructose
49	27	87.1	608	2	C97575	glutamine-fructose
50	27	87.1	608	2	AC2796	hypothetical prote
51	27	87.1	608	2	B93322	Naom Glutamine ami
52	27	87.1	650	2	JC7095	sodium-dependent v
53	27	87.1	661	1	TNBE12	74k alpha trans-in
54	27	87.1	672	2	E70717	hypothetical prote
55	27	87.1	680	2	AH0678	hypothetical prote
56	27	87.1	680	2	A42297	peptidyl carboxy
57	27	87.1	1421	2	T00333	hypothetical prote
58	27	87.1	3228	2	T21381	conserved hypotet
59	26	83.9	80	2	H82194	conserved hypotet
60	26	83.9	102	2	S49794	hypothetical prote
61	26	83.9	114	2	C86401	hypothetical prote
62	26	83.9	121	2	H95121	protein T2205.5 [i
63	26	83.9	174	2	H95121	Tn5252, Orf 9 prot
64	26	83.9	231	2	AB3228	hypothetical prote
65	26	83.9	250	2	C90606	hypothetical prote
66	26	83.9	261	2	T32399	hypothetical prote
67	26	83.9	265	2	T32400	probable hydrolase
68	26	83.9	265	2	T35850	Ctl43 hypothetical
69	26	83.9	268	2	F86522	conserved hypotet
70	26	83.9	268	2	H72101	hypothetical prote
71	26	83.9	287	2	B69148	hypothetical prote
72	26	83.9	297	2	T26464	hypothetical prote
73	26	83.9	322	2	C69528	conserved hypotet
74	26	83.9	325	2	G70771	probable peptidetr
75	26	83.9	331	2	C72015	glycerol-3-phospha
76	26	83.9	331	2	D86610	glycerol-3-P acylt
77	26	83.9	334	2	E90037	conserved hypotet
78	26	83.9	336	2	B84311	histidinol-phospha
79	26	83.9	343	1	A57014	prostatesin (EC 3.4.
80	26	83.9	360	2	T27324	hypothetical prote
81	26	83.9	362	2	C87532	transcription regu
82	26	83.9	431	2	AD1237	arsenic efflux pum
83	26	83.9	454	2	H70158	conserved hypotet
84	26	83.9	468	2	E70839	probable PE protei
85	26	83.9	474	2	T09688	hydroxymethylgluta
86	26	83.9	497	2	D55102	hypothetical prote
87	26	83.9	497	2	F97970	zinc metalloprotei
88	26	83.9	509	2	E90071	hypothetical prote
89	26	83.9	520	2	D90014	hypothetical prote
90	26	83.9	533	2	A44514	anfa protein - Azo
91	26	83.9	549	2	C70689	hypothetical prote
92	26	83.9	571	2	T52576	ATP-dependent RNA
93	26	83.9	641	2	S69651	SAC2 protein - Yea
94	26	83.9	654	1	QRECCS	chemotaxis protein
95	26	83.9	654	2	B85802	hypothetical prote
96	26	83.9	654	2	F90953	sensory transducer
97	26	83.9	703	2	AG0242	probable membrane
98	26	83.9	727	2	T18665	hypothetical prote
99	26	83.9	763	2	T27937	hypothetical prote
100	26	83.9	781	2	H95392	probable MrcB peni
101	26	83.9	812	2	A71335	probable chemotaxi
102	26	83.9	958	2	B83701	alpha-amylase G-6

103	26	83.9	1058	2	A48195	ubiquitin-protein
104	26	83.9	1190	2	F86677	pyruvate-flavodoxi
105	26	83.9	1436	2	F86904	conserved hypochet
106	26	83.9	1626	2	T26318	hypothetical prote
107	26	83.9	1664	2	S67250	DNA-directed RNA p
108	26	83.9	1684	2	S10789	amylose A-180 - al
109	25	80.6	75	2	T03192	hypothetical prote
110	25	80.6	106	2	G90356	hypothetical prote
111	25	80.6	109	2	B95124	conserved hypochet
112	25	80.6	110	2	C86735	conserved hypochet
113	25	80.6	120	2	D97994	conserved hypochet
114	25	80.6	135	2	A24916	ribosomal protein
115	25	80.6	141	2	AE1009	probable Zn(II) - re
116	25	80.6	182	2	AG3449	tripartite transpo
117	25	80.6	185	2	H81837	peptidylprolyl iso
118	25	80.6	185	2	H81103	peptidylprolyl iso
119	25	80.6	186	2	T17497	anhydrin repeat pr
120	25	80.6	201	2	G90084	hypothetical prote
121	25	80.6	205	2	F83333	hypothetical prote
122	25	80.6	208	2	B81327	hypothetical prote
123	25	80.6	212	2	T23211	hypothetical prote
124	25	80.6	218	2	J50163	GTP-binding protei
125	25	80.6	219	2	G84723	probable Ras type
126	25	80.6	219	2	S52024	GTP-binding protei
127	25	80.6	219	2	S52646	GTP-binding protei
128	25	80.6	227	2	C97051	RNA polymerase sig
129	25	80.6	236	2	C35620	coenzyme F420 hydr
130	25	80.6	238	2	A72677	hypothetical prote
131	25	80.6	240	2	B71615	YcpB protein, 1st
132	25	80.6	245	2	T04797	embyo-specific pr
133	25	80.6	248	2	S61960	hypothetical prote
134	25	80.6	251	2	T17397	vir10 protein - Dic
135	25	80.6	257	2	T18969	hypothetical prote
136	25	80.6	270	2	D69791	conserved hypochet
137	25	80.6	293	2	B72325	conserved hypochet
138	25	80.6	297	1	PRSMAG	riboflavin kinase/
139	25	80.6	304	2	A49185	streptogristin A (E
140	25	80.6	315	2	D72086	interphotoreceptor
141	25	80.6	315	2	E86538	conserved hypochet
142	25	80.6	343	2	C75260	CT047 hypochetical
143	25	80.6	352	2	C70532	conserved hypochet
144	25	80.6	359	2	T44332	hypothetical prote
145	25	80.6	360	2	F95933	hypothetical prote
146	25	80.6	362	2	S77677	probable glycosylt
147	25	80.6	362	2	T51464	CTP synthase (EC 6
148	25	80.6	364	2	I45915	RIMC-H2 zinc finger
149	25	80.6	384	2	E83088	interstitial retin
150	25	80.6	398	2	B86298	conserved hypochet
151	25	80.6	398	2	T37083	protein F309.11 i
152	25	80.6	411	1	JC5859	probable transpos
153	25	80.6	414	2	A29835	polyketide synthas
154	25	80.6	430	2	S25977	Sall protein - Esc
155	25	80.6	436	2	F87557	ribosomal protein
156	25	80.6	436	2	G87756	adenylosuccinate l
157	25	80.6	440	2	S49765	protein T01A4.3 i
158	25	80.6	440	2	S69970	TyA protein - yeas
159	25	80.6	440	2	S52895	TyA protein - yeas
160	25	80.6	440	2	F64904	hpa protein - Esc
161	25	80.6	459	2	T29443	probable two-compo
162	25	80.6	465	2	AC0438	conserved hypochet
163	25	80.6	470	2	C95055	6-phospho-beta-gal
164	25	80.6	478	2	H97924	6-phospho-beta-gal
165	25	80.6	478	2	T21353	hypothetical prote
166	25	80.6	492	2	H70959	hypothetical prote
167	25	80.6	511	2	A46020	hypothetical prote
168	25	80.6	511	2	S07095	potassium channel
169	25	80.6	520	2	H83134	potassium channel
170	25	80.6	526	2	T29484	4-Hydroxyphenylace
171	25	80.6	529	2	H81650	hypothetical prote
172	25	80.6	529	2	D71489	conserved hypochet
173	25	80.6	531	2	H72040	hypothetical prote
174	25	80.6	531	2	G86583	conserved hypochet
175	25	80.6	539	2	B72707	CT632 hypochetical
176	25	80.6	554	2	C82700	CTP synthetase XFI
177	25	80.6	556	2	A69718	squalene-hopene cy
178	25	80.6	585	2	A39395	delayed rectifier
179	25	80.6	642	2	A10263	DNA topoisomerase
180	25	80.6	649	2	AG0710	DNA topoisomerase
181	25	80.6	653	2	A85786	DNA topoisomerase
182	25	80.6	653	2	E90937	DNA topoisomerase
183	25	80.6	653	2	JV0049	DNA topoisomerase
184	25	80.6	672	2	T06560	photomorphogenesis
185	25	80.6	675	2	T01112	photomorphogenesis
186	25	80.6	695	2	T40717	hypothetical prote
187	25	80.6	703	2	T04191	hypothetical prote
188	25	80.6	773	2	AE2553	hypothetical prote
189	25	80.6	776	2	S67053	probable membrane
190	25	80.6	781	2	E64222	DNA topoisomerase
191	25	80.6	789	2	S73357	rec
192	25	80.6	831	2	G87620	topoisomerase IV c
193	25	80.6	835	2	S40140	topoisomerase rec
194	25	80.6	837	2	H87638	ribonucleoside-dip
195	25	80.6	840	2	H87639	tonB-dependent rec
196	25	80.6	842	2	AB0674	tonB-dependent rec
197	25	80.6	898	2	A82298	probable hydrolase
198	25	80.6	900	2	JC5224	translacion initia
199	25	80.6	908	2	T22376	methionine-tRNA li
200	25	80.6	923	2	E70820	hypothetical prote
201	25	80.6	925	2	A55748	hypothetical glyci
202	25	80.6	978	2	G75516	protein kinase (EC
203	25	80.6	980	2	T27342	maltooligosyltreha
204	25	80.6	1038	1	WZBE21	hypothetical prote
205	25	80.6	1046	2	T42720	gene 21 protein -
206	25	80.6	1046	2	T42734	cytoplasmic linker
207	25	80.6	1167	2	T06146	cytoplasmic linker
208	25	80.6	1286	1	RJ0BP	disease resistance
209	25	80.6	1307	2	T25563	interphotoreceptor
210	25	80.6	1339	2	A84683	hypothetical prote
211	25	80.6	1392	2	A43336	probable SNF2 subf
212	25	80.6	1427	2	S22635	microtubule-vesicl
213	25	80.6	1592	2	T16055	rescin - human
214	25	80.6	1755	2	S69699	hypochetical prote
215	25	80.6	2133	2	T30637	TYB protein - yeas
216	25	80.6	2391	2	G89779	hypothetical prote
217	25	80.6	2764	2	T13949	hypothetical prote
218	25	80.6	2802	2	T13945	neurofibromin - fr
219	25	80.6	2802	2	T13947	neurofibromin - fr
220	25	80.6	4116	2	T13719	calo protein - fru
221	25	80.6	50	2	E71266	hypochetical prote
222	24	77.4	72	2	JC5628	hypochetical prote
223	24	77.4	99	2	T25759	mandibular organ i
224	24	77.4	107	2	AB3271	hypothetical prote
225	24	77.4	148	2	C82337	hypothetical prote
226	24	77.4	156	2	F91277	hypothetical prote
227	24	77.4	156	2	F68118	hypothetical prote
228	24	77.4	161	2	G75059	hypothetical prote
229	24	77.4	180	2	S60204	hypothetical prote
230	24	77.4	196	2	F97096	trioic protein - co
231	24	77.4	219	2	E90439	ribosomal protein
232	24	77.4	220	2	T24319	conserved hypochet
233	24	77.4	221	2	AB6163	hypochetical prote
234	24	77.4	226	2	G69129	protein F15K9.22 i
235	24	77.4	229	2	AE1300	hypochetical prote
236	24	77.4	229	2	AE1672	ribonuclease III h
237	24	77.4	230	2	F95279	ribonuclease III h
238	24	77.4	253	2	T30928	probable GntR-fam1
239	24	77.4	274	2	T51459	hypothetical prote
240	24	77.4	277	2	S65069	hypothetical prote
241	24	77.4	277	2	AG1037	chitinase (EC 3.2.
242	24	77.4	283	2	B98113	probable capsid pr
243	24	77.4	291	2	T21702	hypothetical prote
244	24	77.4	296	2	I53142	hypothetical prote
245	24	77.4	296	2	A47318	gene Merc protein
246	24	77.4	301	2	C72583	RNA-binding protei
247	24	77.4	315	2	S57696	hypothetical prote
248	24	77.4	323	2	B83464	prohibitin PHB2 -
						flagellar motor sw

249	24	77.4	323	2	B86329	hypothetical prote	322.	24	77.4	889	1	A44153	aconitate hydratase
250	24	77.4	324	2	D64387	hypothetical prote	323.	24	77.4	982	2	T34830	polyketide synthase
251	24	77.4	329	2	A69776	hypothetical prote	324	24	77.4	985	2	S15965	hypothetical prote
252	24	77.4	331	2	A56090	hyaluronoglucuronil	325	24	77.4	1031	2	F83561	probable DNA polym
253	24	77.4	333	2	E95037	carbohydrate kinase	326	24	77.4	1060	2	T43046	retrovirus-related
254	24	77.4	333	2	H97907	2-dehydro-3-deoxyg	327	24	77.4	1060	2	T16507	hypothetical prote
255	24	77.4	336	2	T35215	hypothetical prote	328	24	77.4	1456	1	MMWGPV	RNA-directed RNA p
256	24	77.4	337	2	AD1920	ABC phosphate tran	329	24	77.4	1489	2	G83771	hypothetical prote
257	24	77.4	337	2	H71474	probable peptidogl	330	24	77.4	1638	2	D87749	protein unc-73b [i
258	24	77.4	333	2	G86737	hypothetical prote	331	24	77.4	1846	2	A53289	myr 6, unconventio
259	24	77.4	367	2	A87092	glutamate 5-kinase	332	24	77.4	2139	2	A44467	voltage-dependent
260	24	77.4	369	2	T50660	glutamate 5-kinase	333	24	77.4	2143	2	JH0427	calcium channel pr
261	24	77.4	376	2	E70680	probable prob - My	334	24	77.4	2166	2	S11339	calcium channel al
262	24	77.4	378	2	F91186	probable membrane	335	24	77.4	2220	2	S05054	calcium channel pr
263	24	77.4	378	2	E86033	probable membrane	336	24	77.4	2488	2	A45290	guanine nucleotide
264	24	77.4	378	2	S47807	hypothetical 41.8k	337	24	77.4	2488	2	T42739	hypothetical prote
265	24	77.4	384	2	H96829	probable RNA-bind	338	24	77.4	3429	2	T13853	protein-tyrosine k
266	24	77.4	390	2	T28967	hypothetical prote	339	23	74.2	57	2	PT0201	DNA-directed RNA p
267	24	77.4	401	2	S62465	probable molybdopt	340	23	74.2	60	2	A84270	hypothetical prote
268	24	77.4	408	2	D70365	fimbrial assembly	341	23	74.2	83	2	T07538	heat shock protein
269	24	77.4	410	2	S68188	phosphoglycerate k	342	23	74.2	88	2	PC4238	heat shock protein
270	24	77.4	413	2	S45467	aminopeptidase Y h	343	23	74.2	92	1	S44503	phenol 2-monooxyge
271	24	77.4	447	2	T25264	hypothetical prote	344	23	74.2	94	2	JC1479	heat shock protein
272	24	77.4	459	1	GLAG	beta-glucosidase (345	23	74.2	94	2	A49855	heat shock protein
273	24	77.4	459	1	AI3107	beta-glucosidase (346	23	74.2	94	2	A41884	class I heat-shock
274	24	77.4	466	2	C98179	beta-glucosidase (347	23	74.2	94	2	A83720	chaperonin groES -
275	24	77.4	466	2	T44252	transport protein	348	23	74.2	94	2	S72613	Ig heavy chain V r
276	24	77.4	467	2	A75609	hypothetical prote	349	23	74.2	95	2	S20777	Ig heavy chain V r
277	24	77.4	471	2	D83546	probable amino aci	350	23	74.2	95	2	S24146	S-100 protein p -
278	24	77.4	473	2	B95376	probable ArcD2 arg	351	23	74.2	101	2	AF0223	hypothetical prote
279	24	77.4	474	2	H70896	probable fumarate	352	23	74.2	109	2	D97991	hypothetical prote
280	24	77.4	477	2	G70427	biotin carboxylase	353	23	74.2	109	2	G64609	hypothetical prote
281	24	77.4	478	2	G81679	3-dehydroquininate d	354	23	74.2	116	2	C91279	probable growth in
282	24	77.4	478	2	C71523	probable shikimate	355	23	74.2	116	2	D49339	cell growth regula
283	24	77.4	481	2	B97348	hypothetical prote	356	23	74.2	116	2	D86120	Ig heavy chain V r
284	24	77.4	488	2	B87540	hypothetical prote	357	23	74.2	120	2	A44371	hypothetical prote
285	24	77.4	491	2	D70916	hypothetical glyci	358	23	74.2	126	2	B82448	hypothetical prote
286	24	77.4	494	2	AC0133	probable permease	359	23	74.2	131	2	C65089	hypothetical prote
287	24	77.4	499	2	A47751	NADH2 dehydrogenas	360	23	74.2	132	2	AD0108	conserved hypotet
288	24	77.4	503	2	S14275	steroid 6beta-mono	361	23	74.2	134	2	A88691	protein F4H10.9 [
289	24	77.4	503	2	AH5535	periplasmic dipept	362	23	74.2	135	2	T12863	hypothetical prote
290	24	77.4	506	2	T35261	probable metallope	363	23	74.2	135	2	PC4252	hypothetical 135 p
291	24	77.4	508	2	C82138	conserved hypotet	364	23	74.2	140	2	I38855	centromere protein
292	24	77.4	518	2	D84861	probable pectinest	365	23	74.2	140	2	H71903	hypothetical prote
293	24	77.4	523	2	S53401	probable membrane	366	23	74.2	141	2	A83373	hypothetical prote
294	24	77.4	531	2	B83082	probable binding p	367	23	74.2	142	1	G05630	hypoglycemic horm
295	24	77.4	560	1	RGNVPM	trans-activating t	368	23	74.2	146	2	S85756	partial probable m
296	24	77.4	560	2	T10414	immediate early/pr	369	23	74.2	152	2	S27150	gene 30.3 protein
297	24	77.4	569	2	AG3132	conserved hypotet	370	23	74.2	163	2	S74604	hypothetical prote
298	24	77.4	569	2	E98155	hypothetical prote	371	23	74.2	164	2	T14187	hypothetical prote
299	24	77.4	582	2	B45878	hypothetical prote	372	23	74.2	164	2	AH2183	hypothetical prote
300	24	77.4	589	2	A71277	arginine-tRNA liga	373	23	74.2	164	2	T28044	hypothetical prote
301	24	77.4	592	2	T18243	phenylalanine-tRNA	374	23	74.2	167	2	C97091	nitroreductase fam
302	24	77.4	599	2	T48383	uracil transporter	375	23	74.2	174	2	B72712	hypothetical prote
303	24	77.4	611	2	A48582	vacuolar ATPase A	376	23	74.2	174	2	I49144	ribonucleoprotein
304	24	77.4	611	2	SS6278	DNA-directed DNA p	377	23	74.2	178	2	H64381	cytidylate kinase
305	24	77.4	617	2	T23197	hypothetical prote	378	23	74.2	180	2	A87246	conserved hypotet
306	24	77.4	627	2	AF1074	lytic murein trans	379	23	74.2	180	2	G70912	hypothetical prote
307	24	77.4	631	2	B83404	hypothetical prote	380	23	74.2	182	2	A64007	conserved hypotet
308	24	77.4	643	2	AH0960	cytochrome c-type	381	23	74.2	182	2	AD2914	hypothetical prote
309	24	77.4	643	2	AG0787	cytochrome c-type	382	23	74.2	182	2	G97688	hypothetical prote
310	24	77.4	666	2	E82619	transketonase 1 Xf	383	23	74.2	187	2	E83255	conserved hypotet
311	24	77.4	685	2	F75370	oligopeptidase A -	384	23	74.2	187	2	A99480	hypothetical prote
312	24	77.4	694	2	A95866	probable transketo	385	23	74.2	190	1	A45207	glutathione peroxi
313	24	77.4	707	2	A13290	transketolase (EC	386	23	74.2	194	2	S39798	adenylate kinase (
314	24	77.4	716	2	A75595	hypothetical prote	387	23	74.2	197	2	C81319	hypothetical prote
315	24	77.4	718	2	E71100	hypothetical prote	388	23	74.2	198	2	D75322	conserved hypotet
316	24	77.4	770	2	S60676	cellobiose oxidase	389	23	74.2	208	1	S43851	glutathione transf
317	24	77.4	780	2	T29580	hypothetical prote	390	23	74.2	211	2	A82915	ribosomal protein
318	24	77.4	788	2	B84711	hypothetical prote	391	23	74.2	212	2	G83387	hypothetical prote
319	24	77.4	808	1	P2MWC	2a protein - compe	392	23	74.2	215	2	T22895	hypothetical prote
320	24	77.4	839	2	AC2637	large atp-dependan	393	23	74.2	216	2	H83998	mutants block spor
321	24	77.4	848	2	B97419	probable ATP-depen	394	23	74.2	217	2	F83840	stage II sporulati

395	23	74.2	218	2	B81906	probable membrane
396	23	74.2	222	2	AC2397	ATP-binding protei
397	23	74.2	224	2	T20553	hypothetical prote
398	23	74.2	230	2	T06309	glutathione peroxi
399	23	74.2	231	2	S34448	small nuclear ribo
400	23	74.2	233	2	S39515	H ⁺ -transporting tw
401	23	74.2	234	2	S09760	hypothetical prote
402	23	74.2	235	2	S46590	YnfK protein - yea
403	23	74.2	236	2	G84744	probable calcium-b
404	23	74.2	241	2	A49926	transcription regu
405	23	74.2	245	2	I50511	ribonucleoprotein
406	23	74.2	249	2	T10332	proliferating cell
407	23	74.2	251	2	S59314	hypothetical prote
408	23	74.2	252	2	AH0932	thiamin biosynthes
409	23	74.2	252	2	H97189	glycosyltransferas
410	23	74.2	253	2	T36867	hypothetical prote
411	23	74.2	259	2	B64697	UDP-N-acetylenolpy
412	23	74.2	259	2	B71823	udp-n-acetylenolpy
413	23	74.2	262	2	D82918	conserved hypotet
414	23	74.2	269	2	T32639	hypothetical prote
415	23	74.2	270	2	T40280	26S proteasome reg
416	23	74.2	274	2	G70500	hypothetical prote
417	23	74.2	275	2	F70607	hypothetical prote
418	23	74.2	276	2	T32005	hypothetical prote
419	23	74.2	276	2	G95909	probable 3-deoxy-m
420	23	74.2	280	2	T02004	chitinase (BC 3.2.
421	23	74.2	282	2	AB1929	hypothetical prote
422	23	74.2	285	2	T15498	hypothetical prote
423	23	74.2	285	2	A40657	hypothetical prote
424	23	74.2	288	2	C87342	chemotaxis Mota pr
425	23	74.2	289	2	S54097	Repa protein - Hal
426	23	74.2	291	2	G70605	probable hydrolase
427	23	74.2	293	2	T22385	hypothetical prote
428	23	74.2	294	2	S60991	hypothetical prote
429	23	74.2	294	2	T21668	hypothetical prote
430	23	74.2	295	2	B69514	3-methyladenine DN
431	23	74.2	298	2	D95967	probable transcrip
432	23	74.2	298	2	A75460	conserved hypotet
433	23	74.2	300	2	A84111	sugar transport sy
434	23	74.2	304	2	A83309	hypothetical prote
435	23	74.2	305	2	F71968	hypothetical prote
436	23	74.2	307	2	AG3414	dihydroadipicolinat
437	23	74.2	307	2	T23212	hypothetical prote
438	23	74.2	308	2	AE2359	hypothetical prote
439	23	74.2	309	2	B87457	hypothetical prote
440	23	74.2	310	2	B91247	probable DNA trans
441	23	74.2	310	2	B87667	band 7/Mec-2 famil
442	23	74.2	312	1	S28006	RNA-directed DNA p
443	23	74.2	318	2	T52428	helix-loop-helix p
444	23	74.2	319	2	T37130	probable transport
445	23	74.2	319	2	T36857	conserved hypotet
446	23	74.2	319	2	T46594	phytoene synthase
447	23	74.2	319	2	D90589	hypothetical prote
448	23	74.2	322	2	AF2581	pantothenate kinas
449	23	74.2	323	2	C82234	conserved hypotet
450	23	74.2	323	2	S47966	probable lipid tra
451	23	74.2	325	2	F87049	hypothetical prote
452	23	74.2	327	2	S57121	hypothetical prote
453	23	74.2	328	2	A13511	stomatin like prot
454	23	74.2	329	2	T14760	hypothetical prote
455	23	74.2	330	2	PC4425	lectin-like adhesi
456	23	74.2	333	2	D97363	coaA protein (AE00
457	23	74.2	334	2	AD0775	probable hydrolase
458	23	74.2	334	2	B64977	hypothetical prote
459	23	74.2	334	2	F90991	hypothetical prote
460	23	74.2	334	2	H85836	hypothetical prote
461	23	74.2	338	2	AD0278	probable Lacti-famI
462	23	74.2	343	1	CES0V4	vancomycin resist
463	23	74.2	346	2	A10866	probable oxidoredu
464	23	74.2	346	2	C91090	hypothetical prote
465	23	74.2	346	2	PF5935	hypothetical prote
466	23	74.2	346	2	AD0097	probable aldo/keto
467	23	74.2	346	2	T35363	D-alanine-D-alanin

468	23	74.2	346	2	I39077	RNA-binding protei
469	23	74.2	346	2	C65066	hypothetical prote
470	23	74.2	347	2	A85912	hypothetical prote
471	23	74.2	348	2	B70109	translational releas
472	23	74.2	348	2	B83869	hypothetical prote
473	23	74.2	348	2	B96620	protein T30B16.21
474	23	74.2	348	2	I51677	ribonucleoprotein
475	23	74.2	350	2	AH3020	conserved hypotet
476	23	74.2	350	2	A98264	hypothetical 33.7K
477	23	74.2	350	2	JC2116	hippocampal 38k au
478	23	74.2	351	2	D96761	unknown protein f
479	23	74.2	352	2	G65007	hypothetical prote
480	23	74.2	356	1	S55437	translational releas
481	23	74.2	358	2	G81822	translational releas
482	23	74.2	358	2	G81053	translational releas
483	23	74.2	358	2	H88921	protein F56B10.1 f
484	23	74.2	359	2	T36781	probable glucuronat
485	23	74.2	359	2	I38726	FLAV-like neuronal
486	23	74.2	360	1	B32890	translational releas
487	23	74.2	360	1	FCECR1	translational releas
488	23	74.2	360	2	AB0720	peptide chain rele
489	23	74.2	360	2	B83063	peptide chain rele
490	23	74.2	360	2	A85701	peptide chain rele
491	23	74.2	360	2	AC0246	peptide chain rele
492	23	74.2	360	2	D90843	peptide chain rele
493	23	74.2	360	2	JC6057	RNA-binding protei
494	23	74.2	361	1	I64129	translational releas
495	23	74.2	362	2	AB2713	conserved hypotet
496	23	74.2	362	2	H97494	hypothetical prote
497	23	74.2	363	2	AF2375	hypothetical prote
498	23	74.2	366	2	AB2329	translational releas
499	23	74.2	366	2	I51678	ribonucleoprotein
500	23	74.2	367	2	H82016	translational releas

ALIGNMENTS

RESULT 1
 C64890
 probable electron transfer protein - Escherichia coli (strain K-12)
 C:Species: Escherichia coli
 C>Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 05-Oct-2004
 C:Accession: C64890
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Coi
 .A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A:Title: The complete genome sequence of Escherichia coli K-12.
 A:Reference number: A64720; PMID:97426617; PMID:9278503
 A:Accession: C64890
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-356 <BLAT>
 A:Cross-references: UNIPROT:P76081; UNIPARC:UPI0000131195; GB:AE000236; GB:U00096; NID:g
 A:Experimental source: strain K-12, substrain MG1655
 C:Superfamily: pithalate dioxigenase reductase; cytochrome-b5 reductase homology; ferredox
 C:Keywords: 2Fe-2S; electron transfer; iron-sulfur protein; metalloprotein; transmembrane
 F:116-132/Domain: transmembrane #status predicted <TM01>
 F:284-338/Domain: ferredoxin [2Fe-2S] homology <PFR>
 F:229,304,307,337/Binding site: 2Fe-2S cluster (Cys) (covalent) #status predicted

Query Match 100.0%; Score 31; DB 2; Length 356;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRYAR 6
 |||||
 Db 80 FSRYAR 85

RESULT 2
 C72022
 UDP-N-acetylglucosamine-N-acetyl[muramyl]- (pentapeptide) pyrophosphoryl-undecaprenol N-ac

N/Alternate names: peptidoglycan transferase
C/Species: Chlamydia pneumoniae, Chlamydia pneumoniae
C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C/Accession: C72022; E81518
R/Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A/Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A/Reference number: A72000; PMID:99206606; PMID:10192388
A/Accession: C72022
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-357 <ARN>
A/Cross-references: UNIPROT:Q9Z702; UNIPARC:UPI00000470A7; GB:AE001670; GB:AE001363; NID:
A/Experimental source: strain CML029
R/Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gilm, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A/Title: Genome sequences of Chlamydia trachomatis Mpn and Chlamydia pneumoniae AR39.
A/Reference number: A81500; PMID:20150255; PMID:10684935
A/Accession: E81518
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-357 <REA>
A/Cross-references: UNIPARC:UPI00000470A7; GB:AE002254; GB:AE002161; NID:97189871; PIDN:
A/Experimental source: strain AR39, HL cells
C/Genetics:
A/Gene: murG; CP0962
C/Superfamily: murG protein

Query Match 100.0%; Score 31; DB 2; Length 357;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FSRYAR 6
135 FSRYAR 140

Db 135 FSRYAR 140

RESULT 3
F86603
peptidoglycan transferase [imported] - Chlamydia pneumoniae (strain J138)
C/Species: Chlamydia pneumoniae, Chlamydia pneumoniae
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C/Accession: F86603
R/Shirai, M.; Hiraoka, H.; Kinoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Is
Nucleic Acids Res. 28, 2311-2314, 2000
A/Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A/Reference number: A86491; PMID:20330349; PMID:10871362
A/Accession: F86603
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-357 <STO>
A/Cross-references: UNIPROT:Q9Z702; UNIPARC:UPI00000470A7; GB:BA000008; NID:98979278; PI
A/Experimental source: strain J138
C/Genetics:
A/Gene: murG
C/Superfamily: murG protein

Query Match 100.0%; Score 31; DB 2; Length 357;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FSRYAR 6
135 FSRYAR 140

Db 135 FSRYAR 140

RESULT 4
C69997
probable proline transport protein - Bacillus subtilis
N/Alternate names: probable proline-specific permease
C/Species: Bacillus subtilis
C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 05-Oct-2004

C/Accession: C69997
R/Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Galdwell, B.; Capuano, V.; Carter, N.M.; Choi
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall
leeh, J.; Harwood, C.R.; Hensat, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.;
Koester, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A/Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y., M.; Ogawa, K.; Ogawa, A.; Oudega, B.; Park, S.H.; Paro, V.; Pohl, T.M.; Portetelle,
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,
A/Authors: Schlicht, D.; Schroeder, R.; Scifone, F.; Sekiguchi, J.; Sekowska, A.; Seror,
akushi, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida, K.
A/Authors: Yoshikawa, H.F.; Zumberg, E.; Yoshikawa, H.; Danchin, A.
A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A/Reference number: A69580; PMID:98044033; PMID:9384377
A/Accession: C69997
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-463 <KUN>
A/Cross-references: UNIPROT:Q34618; UNIPARC:UPI0000060942; GB:Z99119; GB:AL009126; NID:97
A/Experimental source: strain 168
C/Genetics:
A/Gene: ynfA
C/Superfamily: ecotropic retrovirus receptor protein
C/Keywords: amino acid transport; transmembrane protein
F/17-33/Domain: transmembrane #status predicted <TM1>
F/44-60/Domain: transmembrane #status predicted <TM2>
F/97-113/Domain: transmembrane #status predicted <TM3>
F/126-142/Domain: transmembrane #status predicted <TM4>
F/158-174/Domain: transmembrane #status predicted <TM5>
F/203-219/Domain: transmembrane #status predicted <TM6>
F/240-256/Domain: transmembrane #status predicted <TM7>
F/280-296/Domain: transmembrane #status predicted <TM8>
F/336-352/Domain: transmembrane #status predicted <TM9>
F/361-377/Domain: transmembrane #status predicted <TM10>
F/404-420/Domain: transmembrane #status predicted <TM11>
F/428-444/Domain: transmembrane #status predicted <TM12>

Query Match 100.0%; Score 31; DB 2; Length 463;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FSRYAR 6
76 FSRYAR 81

Db 76 FSRYAR 81

RESULT 5
H65051
6-phospho-beta-glucosidase (EC 3.2.1.86) - Escherichia coli (strain K-12)
C/Species: Escherichia coli
C/Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C/Accession: H65051; C44070; S27553
R/Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Col
A.; Rose, D.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A/Title: The complete genome sequence of Escherichia coli K-12.
A/Reference number: A64720; PMID:97426617; PMID:9278503
A/Accession: H65051
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-474 <BLAT>
A/Cross-references: UNIPROT:P24240; UNIPARC:UPI00001260EE; GB:AE000355; GB:U00096; NID:97
A/Experimental source: strain K-12, substrain MG1655
R/Hall, B.G.; Xu, L.
Mol. Biol. Evol. 9, 688-706, 1992
A/Title: Nucleotide sequence, function, activation, and evolution of the cryptic asc oper
A/Reference number: A44070; PMID:92334140; PMID:1630307
A/Accession: C44070
A/Status: preliminary
A/Molecule type: DNA

A;Residues: 1-404,'GT','407-427','C','429-454','HR','457-474'<HAL>
A;Cross-references: UNIPARC:UPI000016BEOA; GB:W73326; NID:G145385; PIDN:AAA16430.1; PID:
A;Experimental source: strain IP103
A;Note: sequence extracted from NCBI backbone (NCBIN:109109, NCBIPI:109114)
C;Genetics:
A;Gene: aecB
A;Map position: 59 min
C;Superfamily: Agrobacterium beta-glucosidase
C;Keywords: glycosidase; hydrolase

Query Match 100.0%; Score 31; DB 2; Length 474;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 6; Conservative 0; Mismatches 0; Gaps 0;

QY 1 FSRYPAR 6
|||
157 FSRYPAR 162

Db 157 FSRYPAR 162

RESULT 6
D91075
6-phospho-beta-glucosidase [imported] - Escherichia coli (strain O157:H7, substrain RIMC
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: D91075
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawashi, N.; Yasunaga, T.; Kuhnara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
A;Reference number: A9629; MUID:2156231; PMID:11258796
A;Accession: D91075
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-474 <HAV>
A;Cross-references: UNIPROT:Q8X841; UNIPARC:UPI00000D09A4; GB:BA000007; PIDN:BA836995.1;
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
C;Gene: Ec83572
C;Superfamily: Agrobacterium beta-glucosidase

Query Match 100.0%; Score 31; DB 2; Length 474;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 6; Conservative 0; Mismatches 0; Gaps 0;

QY 1 FSRYPAR 6
|||
157 FSRYPAR 162

Db 157 FSRYPAR 162

RESULT 7
C85920
6-phospho-beta-glucosidase, cryptic [imported] - Escherichia coli (strain O157:H7, subst
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: C85920
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamoultis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: AB5480; MUID:21074935; PMID:11206551
A;Accession: C85920
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-474 <STO>
A;Cross-references: UNIPROT:Q8X841; UNIPARC:UPI00000D09A4; GB:AE005174; NID:G12517164; F
C;Genetics:
A;Gene: aecB
C;Superfamily: Agrobacterium beta-glucosidase

Query Match 100.0%; Score 31; DB 2; Length 474;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 6; Conservative 0; Mismatches 0; Gaps 0;

QY 1 FSRYPAR 6
|||
157 FSRYPAR 162

Db 157 FSRYPAR 162

RESULT 8
B83591
probable transporter PA0443 [imported] - Pseudomonas aeruginosa (strain PA01)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: B83591
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Watterer, P.; Hickey, M.J.; Bri
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lartig, K.; Lim,
; Loay, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathog
A;Reference number: AB2950; MUID:20437337; PMID:10984043
A;Accession: B83591
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-496 <STO>
A;Cross-references: UNIPROT:Q91674; UNIPARC:UPI00000C505A; GB:AE004481; GB:AE004091; NID
C;Genetics:
C;Gene: PA0443
C;Superfamily: Escherichia coli probable transport protein b0511

Query Match 100.0%; Score 31; DB 2; Length 496;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 6; Conservative 0; Mismatches 0; Gaps 0;

QY 1 FSRYPAR 6
|||
257 FSRYPAR 262

Db 257 FSRYPAR 262

RESULT 9
T07092
Ca²⁺-binding EF hand protein homolog PM13 - soybean
C;Species: Glycine max (soybean)
C;Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004
C;Accession: T07092
R;Hsu, T.F.; Tsai, F.-Y.; Heing, Y.I.; Chow, T.Y.
submitted to the EMBL Data Library, May 1997
A;Description: Glycine max mRNA for Ca²⁺-binding EF-hand protein.
A;Reference number: Z15913
A;Accession: T07092
A;Status: preliminary; translated from GB/EMBL/DDBT
A;Molecule type: mRNA
A;Residues: 1-239 <HSU>
A;Cross-references: UNIPROT:Q23959; UNIPARC:UPI00000A8419; EMBL:AF004809; NID:G2270993; f
A;Experimental source: strain Sh1-sh1; cotyledon
C;Genetics:
A;Gene: PM13
C;Superfamily: rice abscisic acid-induced protein

Query Match 90.3%; Score 28; DB 2; Length 239;
Best Local Similarity 83.3%; Pred. No. 47;
Matches 5; Conservative 1; Mismatches 0; Gaps 0;

QY 1 FSRYPAR 6
|||
157 FSRYPAR 162

Db 157 FSRYPAR 162

RESULT 10
T16828
hypothetical protein T07F12.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C;Accession: T16828
R;Chisoe, S.

submitted to the EMBL Data Library, June 1995
 A>Description: The sequence of C. elegans cosmid T07F12.
 A.Reference number: Z18585
 A.Accession: T16828
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A.Molecule type: DNA
 A.Residues: 1-304 <CHI>
 A.Cross-references: UNIPROT:Q22323; UNIPARC:UPI000007A68E; EMBL:U29154; NID:g861403; PID
 A.Experimental source: strain Bristol N2
 C/Genetics:
 A.Gene: CESP:T07F12.1
 A.Introns: 34/2; 74/1; 128/1; 165/1; 208/1

Query Match 90.3%; Score 28; DB 2; Length 304;
 Best Local Similarity 83.3%; Pred. No. 60;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRYAR 6
 Db 147 FARYAR 152

RESULT 11
 AF3271
 probable allantoic permease [imported] - Brucella melitensis (strain 16M)
 C/Species: Brucella melitensis
 C/Date: 01-Feb-2002 #sequence_rev1501 01-Feb-2002 #text_change 09-Jul-2004
 C/Accession: AF3271
 R:DelVecchio, V.G.; Kapur, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
 Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
 A>Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
 A/Reference number: AD3252; PMID:11756688
 A/Accession: AF3271
 A>Status: preliminary
 A.Molecule type: DNA
 A.Residues: 1-409 <KTR>
 A.Cross-references: UNIPROT:Q8YJ01; UNIPARC:UPI0000057B88; GB:AE08917; PIDN:AAU51337.1;
 A.Experimental source: strain 16M
 C/Genetics:
 A.Gene: BME10155
 A/Map position: 1

Query Match 90.3%; Score 28; DB 2; Length 409;
 Best Local Similarity 83.3%; Pred. No. 80;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRYAR 6
 Db 176 FSRYAK 181

RESULT 12
 H95857
 probable deaminase protein [imported] - Sinorhizobium meliloti (strain 1021) megaplasmid
 C/Species: Sinorhizobium meliloti
 C/Date: 24-Aug-2001 #sequence_rev1501 24-Aug-2001 #text_change 31-Dec-2004
 C/Accession: H95857
 R:Finan, T.M.; Weidner, S.; Wong, K.; Bhurmaster, J.; Chain, P.; Vorholter, F.J.; Hernat
 Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
 A>Title: The complete sequence of the 1.683-kb pSymB megaplasmid from the N2-fixing endo
 A/Reference number: A95842; MUID:21396508; PMID:11481431
 A/Accession: H95857
 A>Status: preliminary
 A.Molecule type: DNA
 A.Residues: 1-437 <KTR>
 A.Cross-references: UNIPROT:Q92X34; UNIPARC:UPI00000CB40E; GB:AL591985; PIDN:CAC48528.1;
 A.Experimental source: strain 1021, megaplasmid pSymB
 R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
 J.; Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001
 A/Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kise, E.; Komp, C.; Lelaure,

hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
 A>Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
 A.Reference number: A96039; MUID:21368234; PMID:11474104
 A/Contents: annotation
 C/Genetics:
 A.Gene: Smb20128
 A.Genome: plasmid
 C/Superfamily: metal-dependent hydrolase (amidohydrolase)

Query Match 90.3%; Score 28; DB 2; Length 437;
 Best Local Similarity 83.3%; Pred. No. 85;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRYAR 6
 Db 106 FARYAR 111

RESULT 13
 A95376
 probable ArcD1 arginine/ornithine antiporter [imported] - Sinorhizobium meliloti (strain
 C/Species: Sinorhizobium meliloti
 C/Date: 24-Aug-2001 #sequence_rev1501 24-Aug-2001 #text_change 05-Oct-2004
 C/Accession: A95376
 R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows
 Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
 A>Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti
 A/Reference number: A95262; MUID:21396509; PMID:11481432
 A/Accession: A95376
 A>Status: preliminary
 A.Molecule type: DNA
 A.Residues: 1-475 <KUR>
 A.Cross-references: UNIPROT:Q92Y67; UNIPARC:UPI00000CB248; GB:AE06463; PIDN:AA65571.1;
 A.Experimental source: strain 1021, megaplasmid pSymA
 R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
 J.; Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001
 A/Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kise, E.; Komp, C.; Lelaure,
 hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
 A>Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
 A/Reference number: A96039; MUID:21368234; PMID:11474104
 A/Contents: annotation
 C/Genetics:
 A.Gene: arcD1
 A.Genome: plasmid
 C/Superfamily: ecotropic retrovirus receptor protein

Query Match 90.3%; Score 28; DB 2; Length 475;
 Best Local Similarity 83.3%; Pred. No. 92;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRYAR 6
 Db 220 YSRYAR 225

RESULT 14
 D90202
 methionyl-tRNA synthetase (mets) [imported] - Sulfolobus solfataricus
 C/Species: Sulfolobus solfataricus
 C/Date: 24-May-2001 #sequence_rev1501 24-May-2001 #text_change 09-Jul-2004
 C/Accession: D90202
 R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.U.; Chan-V
 J.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.
 J.; Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001
 A/Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kise, E.; Komp, C.; Lelaure,

A;Residues: 1-573 <KIR>
 A;Cross-references: UNIPROT:Q9UWM2; UNIPARC:UPI00001365BB; GB:AE006641; NID:G13813723; F
 C;Genetics:
 A;Gene: mets
 C;Superfamily: methionyl-tRNA synthetase

Query Match 90.3%; Score 28; DB 2; Length 573;
 Best Local Similarity 83.3%; Pred. No. 1.1e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FSRYPAR 6
 |:||||
 Db 31 FARYPAR 36

RESULT 15
 E64908
 pepidyl-dipeptidase Dcp (EC 3.4.15.5) - Escherichia coli (strain K-12)
 N;Alternate names: dipeptidyl carboxypeptidase II
 C;Species: Escherichia coli
 C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
 C;Accession: E64908; A49931; S14870
 R;Battner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C
 A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A;Title: The complete genome sequence of Escherichia coli K-12.
 A;Reference number: A64720; MUID:97426617; PMID:9278503
 A;Accession: E64908
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-681 <BLAT>
 A;Cross-references: UNIPROT:P24171; UNIPARC:UPI00001680B4; GB:AE000251; GB:U00096; NID:G
 A;Experimental source: strain K-12, substrain MG1655
 R;Henrich, B.; Becker, S.; Schroeder, U.; Plapp, R.
 J. Bacteriol. 175, 7290-7300, 1993
 A;Title: dcp gene of Escherichia coli: cloning, sequencing, transcript mapping, and char
 A;Reference number: A49931; MUID:94042903; PMID:8226676
 A;Accession: A49931
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-138, 'LL', '141-681 <HEN>
 A;Cross-references: UNIPARC:UPI000016F0BD; GB:X57947; NID:G41243; PIDN:CAA41014.1; PID:G
 R;Becker, S.; Plapp, R.
 submitted to the EMBL Data Library, February 1991
 A;Reference number: S14870
 A;Accession: S14870
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-138, 'LL', '141-681 <BEC>
 A;Cross-references: UNIPARC:UPI000016F0BD; EMBL:X57947; NID:G41243; PIDN:CAA41014.1; PID
 C;Genetics:
 A;Gene: dcp
 C;Function:
 A;Note: zinc cofactor
 C;Superfamily: peptidyl-dipeptidase Dcp
 C;Keywords: metalloproteinase; peptidyl-dipeptide hydrolase; zinc

Query Match 90.3%; Score 28; DB 2; Length 681;
 Best Local Similarity 83.3%; Pred. No. 1.3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FSRYPAR 6
 |:||||
 Db 515 FARYPAR 520

RESULT 16
 C90897
 dipeptidyl carboxypeptidase II [imported] - Escherichia coli (strain O157:H7, substrain
 C;Species: Escherichia coli
 C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
 C;Accession: C90897

R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.,
 gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001
 A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genom
 A;Reference number: A99629; MUID:21156231; PMID:11258796
 A;Accession: C90897
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-681 <HAY>
 A;Cross-references: UNIPROT:Q8XB30; UNIPARC:UPI00000D0637; GB:BA000007; PIDN:BA835570.1;
 A;Experimental source: strain O157:H7, substrain RIMD 0509952
 C;Genetics:
 A;Gene: EC92147
 C;Superfamily: peptidyl-dipeptidase Dcp

Query Match 90.3%; Score 28; DB 2; Length 681;
 Best Local Similarity 83.3%; Pred. No. 1.3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FSRYPAR 6
 |:||||
 Db 515 FARYPAR 520

RESULT 17
 D85720
 dipeptidyl carboxypeptidase II [imported] - Escherichia coli (strain O157:H7, substrain I
 C;Species: Escherichia coli
 C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
 C;Accession: D85720
 R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew,
 Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A;Reference number: A85480; MUID:21074935; PMID:11206551
 A;Accession: D85720
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-681 <STO>
 A;Cross-references: UNIPROT:Q8XB30; UNIPARC:UPI00000D0637; GB:AE005174; NID:G12515118; PJ
 A;Experimental source: strain O157:H7, substrain EDL933
 C;Genetics:
 A;Gene: dcp
 C;Superfamily: peptidyl-dipeptidase Dcp

Query Match 90.3%; Score 28; DB 2; Length 681;
 Best Local Similarity 83.3%; Pred. No. 1.3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FSRYPAR 6
 |:||||
 Db 515 FARYPAR 520

RESULT 18
 AC2211
 heme transport protein alx2242 [imported] - Nostoc sp. (strain PCC 7120)
 C;Species: Nostoc sp. PCC 7120
 A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
 C;Accession: AC2211
 R;Kaneh, T.; Nakamura, Y.; Wolz, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriyuchi,
 Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
 DNA Res. 8, 205-213, 2001
 A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anab
 A;Reference number: AB1807; MUID:21595285; PMID:11759840
 A;Accession: AC2211
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-877 <KUR>
 A;Cross-references: UNIPROT:Q8YS49; UNIPARC:UPI00000CE703; GB:BA000019; PIDN:BA874941.1;
 A;Experimental source: strain PCC 7120
 C;Genetics:

A:Gene: alx3242

Query Match 90.3%; Score 28; DB 2; Length 877;
Best Local Similarity 83.3%; Pred. No. 1.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRVAR 6
| | | | |
Db 601 FSRYSR 606

RESULT 19

hypothetical protein PA3728 [imported] - Pseudomonas aeruginosa (strain PA01)
C/Species: Pseudomonas aeruginosa
C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C/Accession: D83181
R/Stover, C.K.; Pham, X.Q.; Ewlin, A.L.; Mizoguchi, S.D.; Warriner, P.; Hickey, M.J.; Bradman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lapidig, K.; Lam, J.; Loty, S.; Olson, M.V.
Nature 406, 959-964, 2000
A/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
A/Reference number: A82950; WUID:20437337; PMID:10984043
A/Accession: D83181
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1746 <STO>
A/Cross-references: UNIPROT:Q9HXR4; UNIPARC:UPI00000C5AE2; GB:AE004791; GB:AE004091; NID
A/Experimental source: strain PA01
C/Genetics:

Query Match 90.3%; Score 28; DB 2; Length 1746;
Best Local Similarity 83.3%; Pred. No. 3.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRVAR 6
| | | | |
Db 1099 FSRVAR 1104

RESULT 20

hypothetical protein At2g37350 [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C/Accession: F84791
R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M. Koo, H.; Mofrac, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Talton, L.; Eues, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A/Reference number: A84420; WUID:20083487; PMID:10617197
A/Accession: F84791
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-36 <STO>
A/Cross-references: UNIPROT:Q9FYA7; UNIPROT:Q9FYB7; UNIPROT:Q8VYA5; UNIPARC:UPI000017A6F
C/Genetics:

A:Gene: At2g37350
A/Map position: 2

Query Match 87.1%; Score 27; DB 2; Length 36;
Best Local Similarity 83.3%; Pred. No. 13;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FSRVAR 6
| | | | |
Db 31 FSRYGR 36

RESULT 21

AI0982

hypothetical protein STY4163 [imported] - Salmonella enterica subsp. enterica serovar Typhimurium
C/Species: Salmonella enterica subsp. enterica serovar Typhimurium
A/Note: this species has also been called Salmonella typhimurium
C/Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C/Accession: AI0982
R/Parthill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Comerford, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moul, S.; O'Garra, P.
Nature 413, 848-852, 2001
A/Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A/Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhimurium
A/Reference number: AB0502; WUID:21534947; PMID:11677608
A/Accession: AI0982
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-43 <PAR>
A/Cross-references: UNIPARC:UPI000005A744; GB:AL513382; PIDN:CAD07989.1; PID:gl6504975; C/Genetics:

A:Gene: STY4163

Query Match 87.1%; Score 27; DB 2; Length 43;
Best Local Similarity 83.3%; Pred. No. 15;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRVAR 6
| | | | |
Db 16 FSRVAR 21

RESULT 22

VC033 protein homolog (AF179595) [imported] - Agrobacterium tumefaciens (strain C58, Cerevisiae)
C/Species: Agrobacterium tumefaciens
C/Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C/Accession: F97521
R/Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Molam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Marx, B.; Science 294, 2323-2328, 2001
A/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens
A/Reference number: A97359; WUID:21608551; PMID:11743194
A/Accession: F97521
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-108 <KUR>
A/Cross-references: UNIPROT:Q8UFR1; UNIPARC:UPI00000D1B08; GB:AE007869; PIDN:AAK87127.1; C/Genetics:

A:Gene: AGR_C_2462
A/Map position: circular chromosome

Query Match 87.1%; Score 27; DB 2; Length 108;
Best Local Similarity 83.3%; Pred. No. 37;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FSRVAR 6
| | | | |
Db 18 FSRVAR 23

RESULT 23

hypothetical protein At2g740 [imported] - Agrobacterium tumefaciens (strain C58)
C/Species: Agrobacterium tumefaciens
C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C/Accession: AH2740
R/Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.; Erge, G.; Giller, W.; Grant, C.; Gunthner, D.; Kuyavin, T.; Levy, R.; Li, M.; McClellan, J.; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, E.; Ster, E.W.
A/Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A/Reference number: AB2577; WUID:21608550; PMID:11743193
A/Accession: AH2740

A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-108 <KUR>
A:Cross-references: UNIPROT:Q8UFRI; UNIPARC:UPI00000DIB08; GB:AE008688; PIDN:AAL42342.1;
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: AFU1336
A:Map position: circular chromosome

Query Match 87.1%; Score 27; DB 2; Length 108;
Best Local Similarity 83.3%; Pred. No. 37;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FSRYPAR 6
Db 18 FERYR 23

RESULT 24

B84920
hypothetical protein At2g47840 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: B84920

R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
Euse, D.; Nierman, W.C.; White, O.; Bisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: B84920
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-208 <STO>
A:Cross-references: UNIPROT:O82251; UNIPARC:UPI00000A1EFL; GB:AE002093; NID:3738296; PI
C:Genetics:
A:Gene: At2g47840
A:Map position: 2

Query Match 87.1%; Score 27; DB 2; Length 208;
Best Local Similarity 83.3%; Pred. No. 71;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FSRYPAR 6
Db 126 FSRYPAR 131

RESULT 25

B49547
nucleoside-diphosphate kinase (EC 2.7.4.6) precursor, mitochondrial - slime mold (Dictyo
C:Species: Dictyostelium discoideum
C>Date: 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change 05-Oct-2004
C:Accession: B49547
R;Troll, H.; Winkler, T.; Lascu, I.; Mueller, N.; Saurin, W.; Veron, M.; Mutzel, R.
J. Biol. Chem. 268, 25469-25475, 1993
A:Title: Separate nuclear genes encode cytosolic and mitochondrial nucleoside diphosphat
A:Reference number: A49547; MUID:94064612; PMID:8244981

A:Accession: B49547
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-220 <TRO>
A:Cross-references: UNIPARC:UPI0000175708; GB:I23068
C:Genetics:
A:Gene: gmk

A:Introns: 25/1; 80/3; 83/1; 173/2
C:Superfamily: nucleoside diphosphate kinase
C:Keywords: ATP binding; hexamer; mitochondrion; phosphonistidine; phosphoprotein; phosp
F/83-87/Region: ATP binding #status predicted
F/186/Active site: His (phosphonistidine intermediate) #status predicted

Query Match 87.1%; Score 27; DB 2; Length 220;
Best Local Similarity 83.3%; Pred. No. 75;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 FSRYPAR 6
Db 2 FSRYPAR 7

RESULT 26

JC4755
ribosomal protein L10.e, cytosolic - fission yeast (Schizosaccharomyces pombe)
N:Alternate names: transcription factor
C:Species: Schizosaccharomyces pombe
C>Date: 10-May-1996 #sequence_revision 16-Aug-1996 #text_change 09-Jul-2004
C:Accession: JC4755; T39755
R;Masson, J.Y.; Vadhais, J.; Ramotar, D.
Gene 170, 153-154, 1996

A:Title: The Schizosaccharomyces pombe spgm gene is a new member of the Qm transcription
A:Reference number: JC4755; MUID:96200877; PMID:8621081
A:Accession: JC4755
A:Molecule type: DNA
A:Residues: 1-232 <MAS>

A:Cross-references: UNIPROT:Q09127; UNIPARC:UPI0000168FD2; GB:U33214; NID:91141785; PIDN:
R;Wood, V.; Barrell, B.G.; Rajandream, M.A.; Xiang, Z.; Aves, S.
submitted to the EMBL Data Library, January 1999
A:Reference number: Z21878
A:Accession: T39755
A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
A:Residues: 1-69, '1', '71-220, 'A' <WOO>
A:Cross-references: UNIPARC:UPI0000133CD5; EMBL:AL035077; PIDN:CAA22664.1; GSPDB:GN00066;
A:Experimental source: strain 972h-; cosmid c18E5
C:Genetics:
A:Gene: spgm; SPDB:SPBC18E5.04

A:Map position: 1
C:Superfamily: ribosomal protein L10
C:Keywords: cytosol; protein biosynthesis; ribosome

Query Match 87.1%; Score 27; DB 2; Length 232;
Best Local Similarity 83.3%; Pred. No. 79;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRYPAR 6
Db 173 FSQYAR 178

RESULT 27

AE3245
hypothetical protein accE [imported] - Agrobacterium tumefaciens (strain C58, Dupont) pl
C:Species: Agrobacterium tumefaciens
C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:Accession: AE3245
R;Wood, D.W.; Serubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.
erage, G.; Gillet, W.; Grant, C.; Guentner, D.; Kutayavin, T.; Levy, R.; Li, M.; McElle
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, I
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB25777; MUID:21608550; PMID:11743193
A:Accession: AE3245
A>Status: preliminary
A:Molecule type: DNA

A:Residues: 1-286 <KUR>
A:Cross-references: UNIPROT:Q8U604; UNIPARC:UPI00000027AF; GB:AE008690; PIDN:AAL46379.1;
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: accE

A:Superfamily: oligopeptide permease protein oppB
A:Genome: plasmid

Query Match 87.1%; Score 27; DB 2; Length 286;
Best Local Similarity 83.3%; Pred. No. 96;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FSRYAR 6
| | | | |
Db 160 FERYAR 165

RESULT 28

T03434
probable transport protein accE - Agrobacterium tumefaciens plasmid pTIC58
C/Species: Agrobacterium tumefaciens
C/Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004
C/Accession: T03434
R.Piper, K.R.; Beck von Bodman, S.; Cook, D.M.; Hwang, I.; Kim, H.; Farrand, S.K.
submitted to the EMBL Data Library, May 1998
A/Reference number: Z14943
A/Accession: T03434
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-286 <PIP>
A/Cross-references: UNIPROT:Q30544; UNIPARC:UPI0000089CE0; EMBL:AF010180; NID:93153171;
C/Genetics:
A/Gene: accE
A/Genome: plasmid pTIC58
C/Superfamily: oligopeptide permease protein oppB

Query Match 87.1%; Score 27; DB 2; Length 286;
Best Local Similarity 83.3%; Pred. No. 96;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FSRYAR 6
| | | | |
Db 160 FERYAR 165

RESULT 29

S35983
proline transport protein - Salmonella typhimurium
C/Species: Salmonella typhimurium
C/Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C/Accession: S35983
R.Liao, M.K.; Maloy, S.
submitted to the EMBL Data Library, August 1993
A/Description: Identification of a cryptic proline transport system in Salmonella typhimurium
A/Reference number: S35983
A/Accession: S35983
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-292 <LIA>
A/Cross-references: UNIPROT:P37460; UNIPARC:UPI000017AA90; EMBL:X74420

Query Match 87.1%; Score 27; DB 2; Length 292;
Best Local Similarity 83.3%; Pred. No. 98;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRYAR 6
| | | | |
Db 96 FSRYAR 101

RESULT 30

E87697
dUDP-4-dehydrohamose reductase [imported] - Caulobacter crescentus
C/Species: Caulobacter crescentus
C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C/Accession: E87697
R.Nierman, W.C.; Paulblum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eissen, J.; Heidelberg, J.B.; Leub, M.T.; Deboy, R.T.; Dodson, R.O.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A/Title: Complete Genome Sequence of Caulobacter crescentus.
A/Reference number: A87249; MUID:21173698; PMID:1159647
A/Accession: E87697

A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-293 <STO>
A/Cross-references: UNIPROT:Q9A2F0; UNIPARC:UPI00000C7B37; GB:AE005673; NID:913425365; P]
C/Superfamily: dUDP-dihydrostrepase synthase

Query Match 87.1%; Score 27; DB 2; Length 293;
Best Local Similarity 83.3%; Pred. No. 99;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FSRYAR 6
| | | | |
Db 150 FSRYAR 155

RESULT 31

T45890
splicing factor-like protein - Arabidopsis thaliana
N/Alternate names: protein PAP12.200
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 05-Oct-2004
C/Accession: T45890
R.Bloeker, H.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Quetier, F.; Salanoubat, M.
submitted to the Protein Sequence Database, January 2000
A/Reference number: Z23016
A/Accession: T45890
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-302 <BLO>
A/Cross-references: UNIPROT:Q9FVB7; UNIPARC:UPI000017A395; EMBL:AL132966
A/Experimental source: cultivar Columbia; BAC clone F4P12
C/Genetics:
A/Map position: 3
A/Intons: 36/2; 41/3; 67/3; 99/3; 164/2; 198/2
A/Note: PAP12.200

Query Match 87.1%; Score 27; DB 2; Length 302;
Best Local Similarity 83.3%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FSRYAR 6
| | | | |
Db 31 FSRYAR 36

RESULT 32

JC5855
polyketide synthase (EC 2.-.-.-) chain 6 - Actinomyadura hibiscia
C/Species: Actinomyadura hibiscia
C/Date: 05-Mar-1998 #sequence_revision 13-Mar-1998 #text_change 09-Jul-2004
C/Accession: JC5855
R.Baird, T.; Hamano, Y.; Igarashi, Y.; Furumai, T.; Oki, T.
Biosci. Biotechnol. Biochem. 61, 1445-1453, 1997
A/Title: Cloning and nucleotide sequence of the putative polyketide synthase genes for p
A/Reference number: JC5850; MUID:97480928; PMID:9339544
A/Accession: JC5855
A/Molecule type: DNA
A/Residues: 1-341 <DAI>
A/Cross-references: UNIPROT:Q32456; UNIPARC:UPI000002F07D; DBJ:D87924; NID:92580441; P1
C/Comment: This enzyme catalyzes repeated condensation cycles of acyl-CoA, affecting the
C/Genetics:
A/Gene: pms6
C/Superfamily: O-methyltransferase
C/Keywords: transferase

Query Match 87.1%; Score 27; DB 2; Length 341;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FSRYAR 6
| | | | |

Db 171 FERRYAR 176

RESULT 33
C83575
fructose-1,6-bisphosphate aldolase PA0555 [imported] - Pseudomonas aeruginosa (strain PA
C/Species: Pseudomonas aeruginosa
C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C/Accession: C83575
R/Stover, C.K.; Pham, X.Q.; Errin, A.L.; Micosguchi, S.D.; Warriner, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
., Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A/Reference number: A82950; MUID:20437337; PMID:10984043
A/Accession: C83575
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-354 <STO>
A/Cross-references: UNIPROT:Q915Y1, UNIPARC:UPI0000125826; GB:AE004492; GB:AE004091; NID:
A/Experimental source: strain PA01
C/Genetics:
A/Gene: fda; PA0555
C/Superfamily: fructose-bisphosphate aldolase II

Query Match 87.1%; Score 27; DB 2; Length 354;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FERRYAR 6
| | | | |
Db 341 FERRYAR 346

RESULT 34
E75080
3-phosphoglycerate kinase (pgk) PAB1679 - Pyrococcus abyssi (strain Orsay)
C/Species: Pyrococcus abyssi
C/Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C/Accession: E75080
R/anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A/Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru
A/Reference number: A75001
A/Accession: E75080
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-410 <KAM>
A/Cross-references: UNIPROT:Q9UZW0, UNIPARC:UPI0000034625; GB:AJ248286; GB:AL096836; NID:
A/Experimental source: strain Orsay
C/Genetics:
A/Gene: PAB1679
C/Superfamily: phosphoglycerate kinase

Query Match 87.1%; Score 27; DB 2; Length 410;
Best Local Similarity 83.3%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FERRYAR 6
| | | | |
Db 91 FERRYAR 96

RESULT 35
D71065
probable phosphoglycerate kinase - Pyrococcus horikoshii
C/Species: Pyrococcus horikoshii
C/Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 16-Aug-2004
C/Accession: D71065
R/Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Seki
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kishida, N.; Oguchi
DNA Res. 5, 55-76, 1998
A/Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a

A/Reference number: A71000; MUID:98344137; PMID:9679194
A/Accession: D71065
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-410 <KAM>
A/Cross-references: UNIPROT:Q58965; UNIPARC:UPI00001180B; GB:AP000005; NID:G3236132; P
A/Experimental source: strain OT3
A/Note: this accession replaces an interim accession for a sequence replaced by GenBank
C/Genetics:
A/Gene: PH1218
C/Superfamily: phosphoglycerate kinase

Query Match 87.1%; Score 27; DB 2; Length 410;
Best Local Similarity 83.3%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FERRYAR 6
| | | | |
Db 91 FERRYAR 96

RESULT 36
T15705
hypothetical protein C30B5.3 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C/Accession: T15705
R/Du, Z.
submitted to the EMBL Data Library, July 1995
A/Description: sequence of C. elegans cosmid C30B5.
A/Reference number: S59417
A/Accession: T15705
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-411 <DUZ>
A/Cross-references: UNIPROT:Q18317, UNIPARC:UPI0000075091; EMBL:U23450; NID:G733552; PID:
A/Experimental source: strain Bristol N2
C/Genetics:
A/Gene: CESP:C30B5.3
A/Introns: 12/2; 91/1; 349/3

Query Match 87.1%; Score 27; DB 2; Length 411;
Best Local Similarity 83.3%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FERRYAR 6
| | | | |
Db 97 FERRYAR 102

RESULT 37
D87672
glutamate-cysteine ligase [imported] - Caulobacter crescentus
C/Species: Caulobacter crescentus
C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 05-Oct-2004
C/Accession: D87672
R/Nierman, W.C.; Feidbylum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DebRoy, R.T.; Dodson, R.J.; Dukkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Emolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A/Title: Complete Genome Sequence of Caulobacter crescentus.
A/Reference number: A87249; MUID:21173698; PMID:11259647
A/Accession: D87672
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-453 <STO>
A/Cross-references: UNIPROT:Q9A2Z2, UNIPARC:UPI00000C7A81; GB:AE005673; NID:G13425126; P
C/Genetics:
A/Gene: CC3414
C/Superfamily: glutamate--cysteine ligase

Query Match 87.1%; Score 27; DB 2; Length 453;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FSRYAR 6
| | | | |
Db 250 FSRYAR 255

RESULT 38

AH0551
proline-specific permease ProY [imported] - *Salmonella enterica* subsp. *enterica* serovar
C/Species: *Salmonella enterica* subsp. *enterica* serovar Typh
C/Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 05-Oct-2004
C/Accession: D90685
R/Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C. G.;
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Rec. 8, 11-22, 2001
A/Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genom
A/Reference number: A99629; MUID:21156231; PMID:11258796
A/Accession: D90685
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-457 <PAR>
A/Cross-references: UNIPARC:UPI000013232F; GB:BA000007; PIDN:BA833875.1;
A/Experimental source: strain O157:H7, substrain RMD 0509952
C/Genetics:
A/Gene: ESO452
C/Superfamily: ecotropic retrovirus receptor protein

Query Match 87.1%; Score 27; DB 2; Length 457;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRYAR 6
| | | | |
Db 75 FSRYAR 80

RESULT 39

B64769
proline transport protein - *Escherichia coli* (strain K-12)
C/Species: *Escherichia coli*
C/Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 05-Oct-2004
C/Accession: B64769
R/Blatner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A/Title: The complete genome sequence of *Escherichia coli* K-12.
A/Reference number: A64720; MUID:97426617; PMID:9278503
A/Accession: B64769
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-457 <BLAT>
A/Cross-references: UNIPROT:P77327; UNIPARC:UPI000013232F; GB:AE000146; GB:U00096; NID:G
A/Experimental source: strain K-12, substrain MG1655
C/Superfamily: ecotropic retrovirus receptor protein
C/Keywords: amino acid transport; transmembrane protein
F/3-59/Domain: transmembrane #status predicted <TM1>
F/96-112/Domain: transmembrane #status predicted <TM2>
F/125-141/Domain: transmembrane #status predicted <TM3>
F/157-173/Domain: transmembrane #status predicted <TM4>
F/241-257/Domain: transmembrane #status predicted <TM5>
F/279-295/Domain: transmembrane #status predicted <TM6>
F/330-346/Domain: transmembrane #status predicted <TM7>
F/357-373/Domain: transmembrane #status predicted <TM8>
F/403-419/Domain: transmembrane #status predicted <TM9>
F/429-445/Domain: transmembrane #status predicted <TM10>

Query Match 87.1%; Score 27; DB 2; Length 457;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRYAR 6
| | | | |
Db 75 FSRYAR 80

RESULT 40

D90685
proline permease transport protein ESO452 [imported] - *Escherichia coli* (strain O157:H7,
C/Species: *Escherichia coli*
C/Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 05-Oct-2004
C/Accession: D90685
R/Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C. G.;
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Rec. 8, 11-22, 2001
A/Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genom
A/Reference number: A99629; MUID:21156231; PMID:11258796
A/Accession: D90685
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-457 <HAY>
A/Cross-references: UNIPROT:P77327; UNIPARC:UPI000013232F; GB:BA000007; PIDN:BA833875.1;
A/Experimental source: strain O157:H7, substrain RMD 0509952
C/Genetics:
A/Gene: ESO452
C/Superfamily: ecotropic retrovirus receptor protein

Query Match 87.1%; Score 27; DB 2; Length 457;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRYAR 6
| | | | |
Db 75 FSRYAR 80

RESULT 41

H85535
proline permease transport protein [imported] - *Escherichia coli* (strain O157:H7, substr
C/Species: *Escherichia coli*
C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 05-Oct-2004
C/Accession: H85535
R/Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew,
Iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A/Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
A/Reference number: A85480; MUID:21074935; PMID:11206551
A/Accession: H85535
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-457 <STO>
A/Cross-references: UNIPROT:P77327; UNIPARC:UPI000013232F; GB:AE005174; NID:G12513248; P
A/Experimental source: strain O157:H7, substrain EDL933
C/Genetics:
A/Gene: proY
C/Superfamily: ecotropic retrovirus receptor protein

Query Match 87.1%; Score 27; DB 2; Length 457;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRYAR 6
| | | | |
Db 75 FSRYAR 80

RESULT 42

AI0388
probable proline-specific permease [imported] - *Yersinia pestis* (strain CO92)
C/Species: *Yersinia pestis*
C/Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 05-Oct-2004
C/Accession: AI0388
R/Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;

deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
 11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
 Nature 413, 523-527, 2001

A>Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.

A/Reference number: AB0001; MUID:21470413; PMID:11586360

A/Accession: A10388

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-463 <KUR>

A/Cross-references: UNIPROT:Q8ZC25; UNIPARC:UPI00000CD9E2; GB:AL590842; PIDN:CAC92436.1;

C/Genetics:

A/Gene: *proX*

C/Superfamily: ecotropic retrovirus receptor protein

Query Match 87.1%; Score 27; DB 2; Length 463;
 Best Local Similarity 83.3%; Pred. No. 1.5e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRYAR 6
 |||||
 Db 77 FSRYAQ 82

RESULT 43

C42603

phospho-beta-glucosidase *ArbB* - *Erwinia chrysanthemi*

C/Species: *Erwinia chrysanthemi*

C/Date: 04-Sep-1998 #sequence_revision 04-Sep-1998 #text_change 09-Jul-2004

C/Accession: C42603

R/El Hassouni, M.; Henriessat, B.; Chipaux, M.; Barrae, F.

J. Bacteriol. 174, 765-777, 1992

A>Title: Nucleotide sequences of the *arb* genes, which control beta-glucoside utilization

in a family including enzymes from eubacteria, archaeobacteria, and humans.

A/Reference number: A42603; MUID:92121114; PMID:1732212

A/Accession: C42603

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-465 <EUL>

A/Cross-references: UNIPROT:P26206; UNIPARC:UPI0000175H1C

A/Note: sequence inconsistent with the nucleotide translation

C/Superfamily: *Agrobacterium* beta-glucosidase

Query Match 87.1%; Score 27; DB 2; Length 465;
 Best Local Similarity 83.3%; Pred. No. 1.5e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FSRYAR 6
 |||||
 Db 150 FERYAR 155

RESULT 44

C87351

aldehyde dehydrogenase [imported] - *Caulobacter crescentus*

C/Species: *Caulobacter crescentus*

C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004

C/Accession: C87351

R/Nierman, W.C.; Felblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

B.; Laub, M.T.; Deboy, R.T.; Dodson, R.O.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A>Title: Complete Genome Sequence of *Caulobacter crescentus*.

A/Reference number: A87249; MUID:21173698; PMID:11259647

A/Accession: C87351

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-478 <STO>

A/Cross-references: UNIPROT:Q9A9Y9; UNIPARC:UPI00000C71A2; GB:AE005673; NID:g13422073; F

C/Genetics:

A/Gene: CC0822

C/Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

Query Match 87.1%; Score 27; DB 2; Length 478;
 Best Local Similarity 83.3%; Pred. No. 1.6e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FSRYAR 6
 |||||
 Db 423 FRYAR 428

RESULT 45

B86565

oligopeptide permease [imported] - *Chlamydia pneumoniae* (strain J138)

C/Species: *Chlamydia pneumoniae*, *Chlamydia pneumoniae*

C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C/Accession: B86565

R/Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ise

Nucleic Acids Res. 28, 2311-2314, 2000

A>Title: Comparison of whole genome sequences of *Chlamydia pneumoniae* J138.

A/Reference number: A86491; MUID:20310349; PMID:10871362

A/Accession: B86565

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-579 <STO>

A/Cross-references: UNIPROT:Q9Z7V7; UNIPARC:UPI0000043CB3; GB:BA000008; NID:98978963; PII

A/Experimental source: strain J138

C/Genetics:

A/Gene: *oppC_2*

Query Match 87.1%; Score 27; DB 2; Length 579;
 Best Local Similarity 83.3%; Pred. No. 1.9e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FSRYAR 6
 |||||
 Db 455 FSRYVR 460

RESULT 46

C72059

peptide ABC transporter, permease protein, probable CP0151 [imported] - *Chlamydia pne*

C/Species: *Chlamydia pneumoniae*, *Chlamydia pneumoniae*

C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004

C/Accession: C72059; B81608

R/Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;

Nature Genet. 21, 385-389, 1999

A>Title: Comparative genomes of *Chlamydia pneumoniae* and *C. trachomatis*.

A/Reference number: A72000; MUID:99206606; PMID:10192388

A/Accession: C72059

A/Molecule type: DNA

A/Residues: 1-579 <ARN>

A/Cross-references: UNIPROT:Q9Z7V7; UNIPARC:UPI0000043CB3; GB:AE001644; GB:AE001363; NID:

A/Experimental source: strain CML029

R/Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, E

, C.; Dodson, R.; Gwinn, M.; Nelson, W.; Deboy, R.; Kolonay, J.; McClarty, G.; Salzberg,

Nucleic Acids Res. 28, 1397-1406, 2000

A>Title: Genome sequences of *Chlamydia trachomatis* MOPn and *Chlamydia pneumoniae* AR39.

A/Reference number: A81500; MUID:20150255; PMID:10684935

A/Accession: B81608

A/Molecule type: DNA

A/Residues: 1-579 <REA>

A/Cross-references: UNIPARC:UPI0000043CB3; GB:AE002176; GB:AE002161; NID:g7189080; PIDN:

A/Experimental source: strain AR39, HL cells

C/Genetics:

A/Gene: *oppC_2*; CP0151

Query Match 87.1%; Score 27; DB 2; Length 579;
 Best Local Similarity 83.3%; Pred. No. 1.9e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FSRYAR 6
 |||||
 Db 455 FSRYVR 460

RESULT 47

SI6561
 glutamine-fructose-6-phosphate transaminase (isomerizing) (EC 2.6.1.16) - Rhizobium meli
 N/Alternate names: glutamine-fructose-6-phosphate aminotransferase
 C/Species: Rhizobium melioli
 C/Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
 C/Accession: SI6561
 R/Author: N.; Endre, G.; Petrovics, G.; Banfalvi, Z.; Kondorosi, A.
 Mol. Gen. Genet. 228, 113-124, 1991
 A/Title: Six nodulation genes of nod box locus 4 in Rhizobium melioli are involved in n
 A/Reference number: SI6561; MUID:91360053; PMID:1909418
 A/Accession: SI6561
 A/Molecule type: DNA
 A/Residues: 1-605 <DB>
 A/Cross-references: UNIPROT:P25195; UNIPARC:UPI000016FEF8; EMBL:X58632; NID:g46331; PIDN
 C/Function:
 A/Description: catalyzes the formation of D-glucosamine-6-phosphate from the amido group
 A/Pathway: glucosamine biosynthesis
 A/Note: glucosamine-6-phosphate is used in the biosynthesis of amino sugars of asparagin
 C/Superfamily: glutamine-fructose-6-phosphate aminotransferase (isomerizing)
 C/Keywords: aminotransferase; isomerase
 F/2-605/Product: glutamine-fructose-6-phosphate transaminase (isomerizing) #status predi
 F/2/Active site: Cys #status predicted

Query Match

Best Local Similarity 87.1%; Score 27; DB 2; Length 605;
 Pred. No. 2e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FSRYAR 6

Db 311 FERYAR 316

RESULT 48

S01040
 glutamine-fructose-6-phosphate transaminase (isomerizing) (EC 2.6.1.16) - Rhizobium legu
 N/Alternate names: nodulation protein nodM
 C/Species: Rhizobium leguminosarum bv. viciae
 C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
 C/Accession: S01040
 R/Surin, B.P.; Downie, J.A.
 Mol. Microbiol. 2, 173-183, 1988
 A/Title: Characterization of the Rhizobium leguminosarum genes nodLMN involved in effici
 A/Reference number: S01039; MUID:88246045; PMID:3132583
 A/Accession: S01040
 A/Molecule type: DNA
 A/Residues: 1-608 <SUR>
 A/Cross-references: UNIPROT:P08633; UNIPARC:UPI000016FEF7B; EMBL:Y00548; NID:g46212; PIDN
 C/Genetics:
 A/Gene: nodM
 A/Genome: plasmid pRLJ1
 C/Superfamily: glutamine-fructose-6-phosphate aminotransferase (isomerizing)
 C/Keywords: aminotransferase; isomerase; nodulation
 F/2-608/Product: glutamine-fructose-6-phosphate transaminase (isomerizing) #status predi
 F/2/Active site: Cys #status predicted

Query Match

Best Local Similarity 87.1%; Score 27; DB 1; Length 608;
 Pred. No. 2e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FSRYAR 6

Db 312 FERYAR 317

RESULT 49

C97575
 glutamine-fructose-6-phosphate transaminase (isomerizing) (EC 2.6.1.16) - Agrobacterium
 N/Alternate names: glucosamine-fructose-6-P aminotransferase (mismomer)
 C/Species: Agrobacterium tumefaciens
 C/Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
 C/Accession: C97575

R.Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
 A.; Liu, F.; Moliam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;
 Science 294, 2323-2328, 2001
 A/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
 A/Reference number: A97359; MUID:21608551; PMID:11743194
 A/Accession: C97575
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-608 <SUR>
 A/Cross-references: UNIPROT:Q8UEH1; UNIPARC:UPI00001643A5; GB:AE007869; PIDN:AAK87556.1;
 C/Genetics:
 A/Gene: AGR C 3284
 A/Map position: circular chromosome
 C/Superfamily: glutamine-fructose-6-phosphate aminotransferase (isomerizing)
 C/Keywords: aminotransferase; isomerase

Query Match

Best Local Similarity 87.1%; Score 27; DB 2; Length 608;
 Pred. No. 2e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FSRYAR 6

Db 312 FERYAR 317

RESULT 50

AC2796
 hypothetical protein glms [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
 C/Species: Agrobacterium tumefaciens
 C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
 C/Accession: AC2796
 R/Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.
 erage, G.; Gallet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClellan
 ; Karp, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, K
 ster, E.W.
 A/Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
 A/Reference number: AB2577; MUID:21608550; PMID:11743193
 A/Accession: AC2796
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-608 <SUR>
 A/Cross-references: UNIPROT:Q8UEH1; UNIPARC:UPI00001643A5; GB:AE008686; PIDN:AA442785.1;
 A/Experimental source: strain C58 (Dupont)
 C/Genetics:
 A/Gene: glms
 A/Map position: circular chromosome
 C/Superfamily: glutamine-fructose-6-phosphate aminotransferase (isomerizing)

Query Match

Best Local Similarity 87.1%; Score 27; DB 2; Length 608;
 Pred. No. 2e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FSRYAR 6

Db 312 FERYAR 317

Search completed: August 29, 2006, 06:14:22
 Job time : 13 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2006, 05:47:27 ; Search time 48 Seconds
(without alignments)
115.627 Million cell updates/sec

Title: US-10-541-343-2

Perfect score: 31

Sequence: 1 PSRYAR 6

Scoring table: BLOSUM62
Gapop.10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database :

1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	31	100.0	324	2	Q2SER3_GGAMM	Q2SER3 habella che
2	31	100.0	338	2	Q7Q9X9_ANOGA	Q7Q9X9 anophelis g
3	31	100.0	356	1	PAAB_ECOLI	P76081 escherichia
4	31	100.0	357	1	MURG_CHLUP	Q92702 chlamydia p
5	31	100.0	412	2	OSMX2_ORYSA	OSMX2 oryza sativ
6	31	100.0	437	2	Q98L45_RHIO	Q98L45 rhizobium l
7	31	100.0	432	2	Q5J1L5_ALCPA	Q5J1L5 alcaligenes
8	31	100.0	447	2	Q62CX5_BURMA	Q62CX5 burkholderi
9	31	100.0	460	2	Q45AY3_9BURK	Q45AY3 burkholderi
10	31	100.0	460	2	Q4L7Z2_9BURK	Q4L7Z2 burkholderi
11	31	100.0	460	2	Q395V1_BURR3	Q395V1 burkholderi
12	31	100.0	461	2	Q2T4Z4_BURR3	Q2T4Z4 burkholderi
13	31	100.0	461	2	Q3JFR3_BURR1	Q3JFR3 burkholderi
14	31	100.0	461	2	Q63M15_BURPS	Q63M15 burkholderi
15	31	100.0	463	2	Q346I8_BACSV	Q346I8 bacillus su
16	31	100.0	468	2	Q2STX5_BURR1	Q2STX5 burkholderi
17	31	100.0	468	2	Q3JY99_BURR1	Q3JY99 burkholderi
18	31	100.0	468	2	Q62G62_BURMA	Q62G62 burkholderi
19	31	100.0	468	2	Q63PW6_BURPS	Q63PW6 burkholderi
20	31	100.0	474	1	ASCB_ECOLI	P24240 escherichia
21	31	100.0	474	2	Q31X80_SHIBS	Q31X80 shigella bo
22	31	100.0	474	2	Q32CL4_SHIBS	Q32CL4 shigella dy
23	31	100.0	474	2	Q3YVE8_SHIBS	Q3YVE8 shigella so
24	31	100.0	474	2	Q2MAB3_ECOLI	Q2MAB3 escherichia
25	31	100.0	474	2	Q8X841_RCO57	Q8X841 escherichia
26	31	100.0	496	2	Q3KAM6_PSEBP	Q3KAM6 pseudomonas
27	31	100.0	496	2	Q2X8L3_PSEPU	Q2X8L3 pseudomonas
28	31	100.0	496	2	Q88FQ2_PSEBP	Q88FQ2 pseudomonas
29	31	100.0	496	2	Q91674_PSEAP	Q91674 pseudomonas
30	31	100.0	526	2	Q3BYW8_XANCS	Q3BYW8 xanthomonas
31	31	100.0	526	2	Q8PQL8_XANNC	Q8PQL8 xanthomonas

32	31	100.0	560	2	Q7SA29_NEUCR	Q7SA29 neurospora
33	31	100.0	635	2	Q21R82_RHOPA	Q21R82 rhodospseudo
34	31	100.0	663	2	Q374J7_RHOPA	Q374J7 rhodospseudo
35	31	100.0	809	2	Q4SCU9_TETNG	Q4SCU9 tetragen n
36	31	100.0	1427	2	Q4P6R1_USTMA	Q4P6R1 ustilago ma
37	31	100.0	1429	2	Q57UM9_9TRYP	Q57UM9 trypanosoma
38	31	100.0	1649	2	Q8MWB3_9TRYP	Q8MWB3 trypanosoma
39	31	100.0	1649	2	Q49100_MYCCA	Q49100 mycoplasma
40	31	100.0	37	2	Q56EC7_9CAUD	Q56EC7 aeromonas p
41	31	100.0	102	2	Q5J3N8_BRARE	Q5J3N8 brachydanio
42	31	100.0	110	2	Q5ZJM2_CHICK	Q5ZJM2 gallus gall
43	31	100.0	133	2	Q3YQB8_BURR1	Q3YQB8 burkholderi
44	31	100.0	170	2	Q52KS6_XENLA	Q52KS6 xenopus lae
45	31	100.0	170	2	Q6DRM0_BRARE	Q6DRM0 brachydanio
46	31	100.0	171	2	Q5AAP5_CANAL	Q5AAP5 candida alb
47	31	100.0	174	1	WRB_HUMAN	WRB_HUMAN
48	31	100.0	174	1	WRB_MOUSE	WRB_MOUSE
49	31	100.0	174	1	WRB_PONPY	WRB_PONPY
50	31	100.0	174	1	WRB_RAT	WRB_RAT
51	31	100.0	174	2	Q3S2Z6_BOVIN	Q3S2Z6 bovis taurus
52	31	100.0	174	2	Q7MXS1_RALEU	Q7MXS1 ralestonia e
53	31	100.0	174	2	Q3TAS5_MOUSE	Q3TAS5 mus musculu
54	31	100.0	174	2	Q52MA2_XENLA	Q52MA2 xenopus lae
55	31	100.0	220	2	Q3EBU3_ARATH	Q3EBU3 arabidopsis
56	31	100.0	227	2	Q3CEN6_THERET	Q3CEN6 thermoaer
57	31	100.0	231	2	Q7XZAI_GRIJA	Q7XZAI griffithsia
58	31	100.0	239	2	Q23959_SORYB	Q23959 glycine max
59	31	100.0	243	2	Q7XQ03_ORYSA	Q7XQ03 oryza sativ
60	31	100.0	243	2	Q9PLN9_ARATH	Q9PLN9 arabidopsis
61	31	100.0	256	2	Q3USC3_BURR1	Q3USC3 burkholderi
62	31	100.0	256	2	Q62KE8_BURMA	Q62KE8 burkholderi
63	31	100.0	256	2	Q63U80_BURPS	Q63U80 burkholderi
64	31	100.0	278	2	Q3N137_9DELT	Q3N137 syntrophoba
65	31	100.0	283	2	Q22323_CABEL	Q22323 caenorhabdi
66	31	100.0	292	2	Q2SW01_BURTH	Q2SW01 burkholderi
67	31	100.0	301	2	Q6UFY8_HORVU	Q6UFY8 hordeum vul
68	31	100.0	306	2	Q4MSJ1_BACCE	Q4MSJ1 bacillus ce
69	31	100.0	306	2	Q63E07_BACCE	Q63E07 bacillus ce
70	31	100.0	306	2	Q6HLI2_BACHK	Q6HLI2 bacillus th
71	31	100.0	306	2	Q81TA2_BACAN	Q81TA2 bacillus an
72	31	100.0	306	2	Q73BD6_BACCI	Q73BD6 bacillus ce
73	31	100.0	322	1	NODZ_RHISN	P55355 rhizobium s
74	31	100.0	322	2	Q85712_RHIFR	Q85712 rhizobium f
75	31	100.0	323	2	Q72VH4_LEPIC	Q72VH4 leprospira
76	31	100.0	323	2	Q8E921_LEPIN	Q8E921 leprospira
77	31	100.0	324	2	Q9AQ17_BRASW	Q9AQ17 bradyrhizob
78	31	100.0	327	2	Q4A5B0_MYC55	Q4A5B0 mycoplasma
79	31	100.0	328	2	Q49HK6_9BACT	Q49HK6 mycoplasma
80	31	100.0	328	2	Q8KUG7_RHIFR	Q8KUG7 rhizobium e
81	31	100.0	329	2	Q8KJ52_RHILIO	Q8KJ52 rhizobium l
82	31	100.0	329	2	Q98AU6_RHILIO	Q98AU6 rhizobium l
83	31	100.0	339	2	Q5X1J2_RAT	Q5X1J2 rattus norv
84	31	100.0	360	2	Q3G6O6_9DELT	Q3G6O6 pelobacter
85	31	100.0	364	2	Q71AA4_9NUCL	Q71AA4 mamestra co
86	31	100.0	364	2	Q8OLC6_9NUCL	Q8OLC6 mamestra co
87	31	100.0	370	1	NODZ_BRAJA	Q45271 bradyrhizob
88	31	100.0	372	2	Q705Z8_CROSA	Q705Z8 crocus sativ
89	31	100.0	377	2	Q6FUJ8_CANGA	Q6FUJ8 candida gla
90	31	100.0	382	2	Q6FUL7_CANGA	Q6FUL7 candida gla
91	31	100.0	397	2	Q92072_RHIME	Q92072 rhizobium m
92	31	100.0	399	2	Q4CH65_CLOTHM	Q4CH65 clostridium
93	31	100.0	409	2	Q8YD11_BRUME	Q8YD11 bruceella me
94	31	100.0	420	2	Q3CKC0_THERET	Q3CKC0 thermoaer
95	31	100.0	437	2	Q92X34_RHIME	Q92X34 rhizobium m
96	31	100.0	445	1	FDPT_VARI1	Q9Y753 yarrowia li
97	31	100.0	448	2	Q41YG5_AZOVIT	Q41YG5 azotobacter
98	31	100.0	468	2	Q3SIH2_THIDA	Q3SIH2 thiodiazolu
99	31	100.0	472	2	Q3JAX0_NITOC	Q3JAX0 nitrosococ
100	31	100.0	473	2	Q6AC11_LEIIX	Q6AC11 leifeonia x
101	31	100.0	475	2	Q92Y77_RHIME	Q92Y77 rhizobium m
102	31	100.0	476	2	Q61V14_PECOC	Q61V14 peccobacter
103	31	100.0	476	2	Q6D574_ERWCT	Q6D574 erwina car
104	31	100.0	479	2	Q89DW4_BRAJA	Q89DW4 bradyrhizob

105	28	90.3	486	2	Q40P9_9BURK	Q4aqp9	polaromonas
106	28	90.3	487	2	Q6BMH4_DEBHA	Q6bmh4	debaromyce
107	28	90.3	487	2	Q8FYG8_BRUBA	Q8fyg8	bruceella su
108	28	90.3	495	2	Q4KDW7_PSEF5	Q4kdw7	pseudomonas
109	28	90.3	496	2	Q3H3P9_9ACTO	Q3h3p9	nocardioide
110	28	90.3	502	1	AMPA_DESPS	Q6ajz2	desulfocale
111	28	90.3	538	2	Q6ESG6_ORYSA	Q6esg6	oryza sativ
112	28	90.3	548	2	Q4WCP9_ASPFU	Q4wcp9	aspergillus
113	28	90.3	548	2	Q6MEX0_PARUW	Q6mex0	parachlamyd
114	28	90.3	571	1	SYM_SULAC	Q4jhm3	sulfolobus
115	28	90.3	571	1	SYM_SULSO	Q57ic1	sulfolobus
116	28	90.3	573	1	SYM_SULSO	Q5uwr2	sulfolobus
117	28	90.3	576	2	Q9FD23_PPSED	Q9fd23	pseudomonas
118	28	90.3	582	2	Q8H399_ORYSA	Q8h399	oryza sativ
119	28	90.3	605	2	Q2UUI2_ASPOR	Q2uii2	aspergillus
120	28	90.3	658	2	Q2MAB7_MAGSA	Q2mab7	magnetospir
121	28	90.3	680	1	DCP_BCOLI	P24171	escherichia
122	28	90.3	680	2	Q4O3H9_LEIMA	Q4g3h9	leishmania
123	28	90.3	681	2	Q32ON6_SHIBS	Q32on6	shigella bo
124	28	90.3	681	2	Q32G39_SHIDS	Q32g39	shigella dy
125	28	90.3	681	2	Q32IS4_SHISS	Q32is4	shigella so
126	28	90.3	681	2	Q7UCH1_SHIFL	Q7uch1	shigella fl
127	28	90.3	681	2	Q83IL2_SHIFL	Q83il2	shigella fl
128	28	90.3	681	2	Q8XB30_TCO57	Q8xb30	escherichia
129	28	90.3	686	2	Q4SOS9_TETNG	Q4sgs9	tetrarodon n
130	28	90.3	703	2	Q8FHD3_ECOLI	Q8fhd3	escherichia
131	28	90.3	758	2	Q4BD58_BURVI	Q4bd58	burkholderi
132	28	90.3	766	2	Q31N21_SYNP7	Q31n21	synecococc
133	28	90.3	766	2	Q5M227_SYNP6	Q5m227	synecococc
134	28	90.3	822	2	Q5LIG5_BACFN	Q5liq5	bacteroides
135	28	90.3	877	2	Q8Y849_ANASP	Q8y849	anaeana sp
136	28	90.3	878	2	Q38BP9_9TRYP	Q38bp9	trypanosoma
137	28	90.3	898	2	Q6MT97_MYCCA	Q6mt97	mycoplasma
138	28	90.3	899	2	Q2SSJ3_MYCCA	Q2ssj3	mycoplasma
139	28	90.3	927	2	Q4WML1_ASPFU	Q4wml1	aspergillus
140	28	90.3	936	2	Q9DMD4_RCMNM	Q9dmd4	rat cytoleg
141	28	90.3	970	2	Q91TM6_TUHYV	Q91tm6	tupaiid her
142	28	90.3	993	2	Q4WIN7_ASPFU	Q4win7	aspergillus
143	28	90.3	1005	2	Q7NNY2_GLOVI	Q7nn2	gloeobacter
144	28	90.3	1064	2	Q74FA0_GEOSL	Q74fa0	gloeobacter s
145	28	90.3	1069	2	Q472S2_GOLP1	Q472s2	colwellia p
146	28	90.3	1206	2	Q5F9S0_NIEGI	Q5f9s0	nitrospira g
147	28	90.3	1242	2	Q4SS72_TETNG	Q4ss72	tetrarodon n
148	28	90.3	1746	2	Q9HXK4_PSEAE	Q9hxk4	pseudomonas
149	28	90.3	2168	2	Q90ZAF_BRARE	Q90zaf	brachydanio
150	28	90.3	2196	2	Q5T2F1_BRARE	Q5t2f1	brachydanio
151	28	90.3	2289	2	Q2LYX1_DROPS	Q2lyx1	drosochilla
152	28	90.3	2401	2	Q7RF52_PLAYO	Q7rf52	plasmodium
153	27	87.1	443	2	Q8Z2A2_SALTI	Q8z2a2	salmonella
154	27	87.1	102	2	Q6U965_SCAUD	Q6u965	aeromonas p
155	27	87.1	103	2	Q2LYE1_9DELT	Q2lye1	syntrichophus
156	27	87.1	107	2	Q5LMH4_SILPO	Q5lmh4	silicibacte
157	27	87.1	108	2	Q8UFM1_AGRIS	Q8ufm1	agrobacteri
158	27	87.1	126	2	Q424A3_DBSHA	Q424a3	desulfiflaba
159	27	87.1	128	2	Q43ZB5_SOLUS	Q43zb5	solibacter
160	27	87.1	134	2	Q4VXS1_VIGMU	Q4vxs1	vigna mungo
161	27	87.1	137	2	Q2QZN6_ORYSA	Q2qzn6	oryza sativ
162	27	87.1	145	2	Q8MWE0_BENMO	Q8mwe0	penaeus mon
163	27	87.1	153	2	Q3W196_9ACTO	Q3w196	frankia sp.
164	27	87.1	171	2	Q6DU58_THELAN	Q6du58	helianthus
165	27	87.1	175	2	Q8ZVAS_NITBU	Q8zvas	nitrosomona
166	27	87.1	183	2	Q813H6_PLAUF	Q813h6	plasmodium
167	27	87.1	186	2	Q21XV0_RHOPA	Q21xv0	rhodopseud
168	27	87.1	208	2	Q82251_ARATH	Q82251	arabidopsis
169	27	87.1	208	2	Q94A61_ARATH	Q94a61	arabidopsis
170	27	87.1	209	2	Q8LBB6_ARATH	Q8lbb6	arabidopsis
171	27	87.1	209	2	Q9FM67_ARATH	Q9fm67	arabidopsis
172	27	87.1	219	2	Q2VZX2_MAGSA	Q2vzx2	magnetospir
173	27	87.1	220	1	NDKM_DICDI	P34093	dictyosteli
174	27	87.1	220	1	Q54WJ3_DICDI	Q54wj3	dictyosteli
175	27	87.1	221	1	RL10A_SCHPO	009127	schistosach
176	27	87.1	221	1	RL10B_SCHPO	Q9p766	schistosach
177	27	87.1	222	2	Q5L844_BACFN	Q5l844	bacterioides

251	27	87.1	453	2	06TM91_9POXV	06TM91_orf_virus.	324	27	87.1	581	2	Q2IEU8_9DELT	Q2IEU8_aeaeomyxob
252	27	87.1	455	2	05XKG4_GEOKA	05XKG4_geobacillus	325	27	87.1	583	2	Q3B612_PELLD	Q3B612_peloidictyon
253	27	87.1	456	1	PROT_SALTU	P37460_salmoneila	326	27	87.1	589	2	081013_CAEBL	081013_caenorthabdi
254	27	87.1	456	2	05PPT1_SALPA	Q5PT1_salmoneila	327	27	87.1	592	1	S23A2_RAT	Q9WU8_rattus_nov
255	27	87.1	456	2	08ZBY3_SALTU	P0AEE3_salmoneila	328	27	87.1	594	1	CPB2_CAEDA	Q6E3F0_caenorthabdi
256	27	87.1	457	1	PROX_EGOS7	P0AEE3_escherichia	329	27	87.1	601	2	Q7NTV8_CHRVO	Q7NTV8_chromobacte
257	27	87.1	457	1	PROX_ECOL6	P0AEE3_escherichia	330	27	87.1	604	1	NODM2_RHIME	P25195_rhizobium_m
258	27	87.1	457	1	PROX_ECOL6	P0AEE3_escherichia	331	27	87.1	606	1	GLMS_RHITO	Q981X5_r_glucoam
259	27	87.1	457	1	PROX_ECOL6	P0AEE3_escherichia	332	27	87.1	607	1	GLMS_AGRTS	Q981X5_r_glucoam
260	27	87.1	457	2	Q3Z507_SHISS	Q3Z507_shigella_so	333	27	87.1	607	1	GLMS_RHIME	Q92B54_r_glucoam
261	27	87.1	457	2	Q2NNA54_9SPHN	Q2NNA54_erythrobact	334	27	87.1	607	1	NODM1_RHIME	Q92B54_rhizobium_m
262	27	87.1	457	2	Q2MC24_ECOLI	Q2MC24_escherichia	335	27	87.1	607	1	NODM1_RHITV	Q08633_rhizobium_1
263	27	87.1	457	2	Q3Z508_SHISS	Q3Z508_shigella_bo	336	27	87.1	607	2	Q2UPA8_ASPOR	Q2UPA8_aspergillus
264	27	87.1	457	2	Q3ZJF4_SHIDS	Q3ZJF4_shigella_dy	337	27	87.1	607	2	Q8KJME_RHITO	Q8KJME_rhizobium_1
265	27	87.1	457	2	Q83SG9_SHIFL	Q83SG9_shigella_fl	338	27	87.1	607	2	Q988E3_RHITO	Q988E3_rhizobium_1
266	27	87.1	458	2	Q6D860_ERWCT	Q6D860_erynia_car	339	27	87.1	607	2	Q989K4_RHITO	Q989K4_rhizobium_1
267	27	87.1	463	2	Q8D167_YERPE	Q8D167_yersinia_pe	340	27	87.1	608	2	Q2XRG1_RHLET	Q2XRG1_rhizobium_e
268	27	87.1	463	2	Q66DX2_YERPS	Q66DX2_yersinia_ps	341	27	87.1	610	2	Q7RD29_YVTRU	Q7RD29_oryzster_mush
269	27	87.1	463	2	Q7NYY4_CHRVO	Q7NYY4_chromobacte	342	27	87.1	615	2	Q706Q0_PSEPU	Q706Q0_pseudomonas
270	27	87.1	463	2	Q8ZC25_YERPE	Q8ZC25_yersinia_pe	343	27	87.1	619	2	Q7JM92_CAEBL	Q7JM92_caenorthabdi
271	27	87.1	464	1	CASB_KTEOX	Q48409_klebsiella	344	27	87.1	620	2	Q6Z8B3_CAEBR	Q6Z8B3_caenorthabdi
272	27	87.1	464	2	Q8KP24_KLEAE	Q8KP24_klebsiella	345	27	87.1	620	2	Q2UPV4_ASPOR	Q2UPV4_aspergillus
273	27	87.1	464	1	ARB8_ERMCH	Q6D517_erynia_car	346	27	87.1	629	2	Q8NNV5_CORGL	Q8NNV5_corynebacte
274	27	87.1	465	1	ARB8_ERMCH	P26306_erynia_car	347	27	87.1	630	2	Q7N903_PHOLL	Q7N903_photobactdu
275	27	87.1	467	2	Q9GPIJ_LEIIN	Q9GPIJ_leishmania	348	27	87.1	638	2	Q67QV8_SYWTH	Q67QV8_symbiobacte
276	27	87.1	468	2	Q6QGV5_PECOC	Q6QGV5_pseudobacter	349	27	87.1	642	2	S23A2_MOUSE	Q9EPT4_mus_musculu
277	27	87.1	472	2	Q2XCE9_PSEBU	Q2XCE9_pseudomonas	350	27	87.1	648	1	Q80Y23_MOUSE	Q80Y23_mus_musculu
278	27	87.1	472	2	Q88AN24_PSEBK	Q88AN24_pseudomonas	351	27	87.1	650	1	S23A2_HUMAN	Q9UW3_homo_sapien
279	27	87.1	473	2	Q9Z664_ZYMMO	Q9Z664_zymomonas_m	352	27	87.1	651	2	Q9Y576_PIG	Q9Y576_sus_scrofa
280	27	87.1	473	2	Q3K7F6_PSEPF	Q3K7F6_pseudomonas	353	27	87.1	660	1	AT12_VZVD	Q74421_myobacteri
281	27	87.1	473	2	Q4K6Z6_PSEPF	Q4K6Z6_pseudomonas	354	27	87.1	661	2	Q2B152_HMY3	Q2B152_human herpes
282	27	87.1	475	2	Q5N1M2_ZYMMO	Q5N1M2_zymomonas_m	355	27	87.1	661	2	Q74421_MYCPA	Q74421_myobacteri
283	27	87.1	478	2	Q5BPM0_EMENT	Q5BPM0_aspergillus	356	27	87.1	661	2	Q6QCP3_HMY3	Q6QCP3_human herpes
284	27	87.1	481	2	Q9A9Y9_CAUCR	Q9A9Y9_caulobacter	357	27	87.1	672	1	Y959_MYCTU	Q6A5D7_myobacteri
285	27	87.1	483	2	Q48G57_PSE14	Q48G57_pseudomonas	358	27	87.1	672	1	Y984_MYCBO	Q9B591_homo_sapien
286	27	87.1	483	2	Q4ZOK9_PSEU2	Q4ZOK9_pseudomonas	359	27	87.1	673	2	Q9BT91_HUMAN	Q9BT91_homo_sapien
287	27	87.1	483	2	Q885G3_PSEBM	Q885G3_pseudomonas	360	27	87.1	675	2	Q58EB80_XENIA	Q58EB80_xenopus_lae
288	27	87.1	484	2	Q7Z9M2_TTRRE	Q7Z9M2_trichoderma	361	27	87.1	680	1	DCP_SALTU	P27236_salmoneila
289	27	87.1	488	2	Q9AEN6_ACTNA	Q9AEN6_actinomyces	362	27	87.1	680	2	Q57FC6_SALCH	Q57FC6_salmoneila
290	27	87.1	488	2	Q4SUT2_TETNG	Q4SUT2_tetradon_n	363	27	87.1	680	2	Q5PHG4_SALPA	Q5PHG4_salmoneila
291	27	87.1	492	2	Q39B85_BURB3	Q39B85_burkholderi	364	27	87.1	680	2	Q8Z627_SALTI	Q8Z627_salmoneila
292	27	87.1	494	2	Q3F322_9BURK	Q3F322_burkholderi	365	27	87.1	687	2	Q346T5_RHOPA	Q346T5_rhodopseudo
293	27	87.1	494	2	Q450X1_9BURK	Q450X1_burkholderi	366	27	87.1	687	2	Q35N66_9BRAD	Q35N66_bradynrhizob
294	27	87.1	494	2	Q4B8B6_BURVI	Q4B8B6_burkholderi	367	27	87.1	687	2	Q35YR7_RHOPA	Q35YR7_rhodopseudo
295	27	87.1	494	2	Q4BTR1_BURVI	Q4BTR1_burkholderi	368	27	87.1	687	2	Q37L11_RHOPA	Q37L11_rhodopseudo
296	27	87.1	494	2	Q4LMW0_9BURK	Q4LMW0_burkholderi	369	27	87.1	687	2	Q3F2C5_NITWA	Q3F2C5_nitrobacter
297	27	87.1	494	2	Q39F92_BURB3	Q39F92_burkholderi	370	27	87.1	687	2	Q3SPA9_NITWA	Q3SPA9_nitrobacter
298	27	87.1	495	2	Q45A58_9BURK	Q45A58_burkholderi	371	27	87.1	687	2	Q3UG96_BURP1	Q3UG96_burkholderi
299	27	87.1	495	2	Q4LQW2_9BURK	Q4LQW2_burkholderi	372	27	87.1	687	2	Q2J0W1_RHOPA	Q2J0W1_rhodopseudo
300	27	87.1	496	2	Q2SWW0_BURTH	Q2SWW0_burkholderi	373	27	87.1	687	2	Q6NNA3_RHOPA	Q6NNA3_rhodopseudo
301	27	87.1	497	2	Q6MYG5_ASPFU	Q6MYG5_aspergillus	374	27	87.1	687	2	Q8S5D8_BRADA	Q8S5D8_bradynrhizob
302	27	87.1	497	2	Q4WRG4_ASPFU	Q4WRG4_aspergillus	375	27	87.1	717	2	Q4SDQ2_TERNG	Q4SDQ2_tetradon_n
303	27	87.1	500	2	Q88Y80_LACPL	Q88Y80_lactobacill	376	27	87.1	726	2	Q3JPV5_BURP1	Q3JPV5_burkholderi
304	27	87.1	502	2	Q82GV1_STRPM	Q82GV1_streptomyce	377	27	87.1	735	2	Q3GLZ1_9GAMM	Q3GLZ1_psychrobact
305	27	87.1	510	2	Q57SG3_SALCH	Q57SG3_salmoneila	378	27	87.1	737	1	AMY1_AEDAE	P53354_aedes_aegyp
306	27	87.1	513	2	Q3F681_9BURK	Q3F681_burkholderi	379	27	87.1	737	2	Q02413_AEDAE	Q02413_aedes_aegyp
307	27	87.1	513	2	Q87143_VIBPA	Q87143_vibrio_para	380	27	87.1	738	2	Q4B2Y7_9BURK	Q4B2Y7_polatromas
308	27	87.1	525	2	Q3WZEB_9ACTN	Q3WZEB_rubrobacter	381	27	87.1	752	2	Q3PUT5_PARDE	Q3PUT5_pateococcus
309	27	87.1	530	2	Q5HKM4_STABO	Q5HKM4_staphylococ	382	27	87.1	761	2	Q34DB7_RHOPA	Q34DB7_rhodopseudo
310	27	87.1	530	2	Q8CQ99_STARS	Q8CQ99_staphylococ	383	27	87.1	762	2	Q37FJ0_RHOPA	Q37FJ0_rhodopseudo
311	27	87.1	536	2	Q873X2_FUSSP	Q873X2_fusarium_sp	384	27	87.1	762	2	Q2IMW27_RHOPA	Q2IMW27_rhodopseudo
312	27	87.1	542	2	Q8U1Z7_FUSSP	Q8U1Z7_fusarium_sp	385	27	87.1	763	2	Q36VU4_RHOPA	Q36VU4_rhodopseudo
313	27	87.1	542	2	Q96VQ9_FUSSP	Q96VQ9_fusarium_sp	386	27	87.1	763	2	Q6N6M1_RHOPA	Q6N6M1_rhodopseudo
314	27	87.1	543	2	Q810P0_CAEBL	Q810P0_caenorthabdi	387	27	87.1	764	2	Q3PU98_NITWA	Q3PU98_nitrobacter
315	27	87.1	544	2	Q4UCP8_THEAN	Q4UCP8_theileria_a	388	27	87.1	766	2	Q3SSN4_NITWA	Q3SSN4_nitrobacter
316	27	87.1	547	2	Q3PVG4_NITWA	Q3PVG4_nitrobacter	389	27	87.1	784	2	Q3SEBP6_9BRAD	Q3SEBP6_bradynrhizob
317	27	87.1	549	2	Q5XGN4_XENIA	Q5XGN4_xenopus_lae	390	27	87.1	799	2	Q6ZCR4_ORYSA	Q6ZCR4_oryza_sativ
318	27	87.1	557	2	Q4NSC9_ASPFU	Q4NSC9_aspergillus	391	27	87.1	848	2	Q6MPE2_BDRBA	Q6MPE2_bdellovibri
319	27	87.1	560	2	Q2UCU1_ASPOR	Q2UCU1_aspergillus	392	27	87.1	848	2	Q891Q4_BRADA	Q891Q4_bradynrhizob
320	27	87.1	566	1	CPB2_CAEBR	Q6E3F3_caenorthabdi	393	27	87.1	899	2	Q3XJ30_9PROT	Q3XJ30_magnetococ
321	27	87.1	571	1	CPB2_CAEBR	Q6E3F2_caenorthabdi	394	27	87.1	963	2	Q5S071_THET8	Q5S071_thermus the
322	27	87.1	575	2	Q6ZRB0_HUMAN	Q6ZRB0_homo_sapien	395	27	87.1	963	2	Q72UJ4_THET8	Q72UJ4_thermus the
323	27	87.1	579	2	Q9Z7V7_CHLBN	Q9Z7V7_chlamydia_p	396	27	87.1	1014	2	Q2IPV0_9DELT	Q2IPV0_aeaeomyxob

```

397 27 87.1 1130 2 Q4SV4_TETNG
398 27 87.1 1170 2 Q2T4P2_BURTH
399 27 87.1 1240 2 Q38B4_9TRYR
400 27 87.1 1377 2 Q9P2A8_HUMAN
401 27 87.1 1467 2 Q4RMJ3_TETNG
402 27 87.1 1481 2 Q9U107_CABEL
403 27 87.1 1481 2 Q8CFQ3_MOUSE
404 27 87.1 1481 2 Q3ULE8_MOUSE
405 27 87.1 1500 2 Q80TX8_MOUSE
406 27 87.1 1521 2 Q60306_HUMAN
407 27 87.1 1542 2 Q2SHZ2_9GAWM
408 27 87.1 1632 2 Q93593_CABEL
409 27 87.1 1636 2 Q40Y08_KINRA
410 27 87.1 1833 2 Q570B6_9TRYR
411 27 87.1 1954 2 Q4WKS1_ASPFU
412 27 87.1 2006 2 Q5BGD9_EMENT
413 27 87.1 2089 2 Q4D8B9_9TRYR
414 27 87.1 2126 2 Q6P4S8_MOUSE
415 27 87.1 2798 2 Q4DMW7_9TRYR
416 27 87.1 2982 2 Q4RXP0_TETNG
417 27 87.1 3167 2 Q7Q087_GIALA
418 26 83.9 68 2 Q8R728_THETN
419 26 83.9 77 2 Q2VZM8_MAGSA
420 26 83.9 80 2 Q9KRZ6_VIBCH
421 26 83.9 84 2 Q5BT57_SCHUA
422 26 83.9 98 2 Q58PS6_9PROT
423 26 83.9 102 1 Y116_YEAST
424 26 83.9 102 2 Q7W3B4_BORPA
425 26 83.9 105 2 Q3AM77_SYNSC
426 26 83.9 105 2 Q44SD5_CHILI
427 26 83.9 107 2 Q6Z565_ORYSA
428 26 83.9 107 2 Q9R6Q7_PROMI
429 26 83.9 112 2 Q3X4U8_9ACTN
430 26 83.9 114 2 Q9SFEZ_ARATH
431 26 83.9 116 2 Q3RRUS_RALME
432 26 83.9 116 2 Q4N1J9_9M1CC
433 26 83.9 119 2 Q2XVU1_9CAUD
434 26 83.9 119 2 Q3HLE6_9CAUD
435 26 83.9 119 2 Q2LIC3_9CAUD
436 26 83.9 119 2 Q2L1H6_9CAUD
437 26 83.9 119 2 Q3H1I0_9CAUD
438 26 83.9 119 2 Q218P3_9CAUD
439 26 83.9 121 2 Q3D1N5_STRAG
440 26 83.9 121 2 Q3DLF5_STRAG
441 26 83.9 121 2 Q548B4_STRPV
442 26 83.9 121 2 Q8DZ66_STRAS
443 26 83.9 121 2 Q8E4R2_STRAR
444 26 83.9 121 2 Q970Y7_STRPV
445 26 83.9 131 2 Q3VL18_9CHLB
446 26 83.9 131 2 Q3ASB5_CHLCH
447 26 83.9 136 2 Q43GL7_9CHLB
448 26 83.9 138 2 Q3B6N1_PELLD
449 26 83.9 139 2 Q3AHZ9_9CHLB
450 26 83.9 147 2 Q8U1A4_9CAUD
451 26 83.9 151 2 Q2XG16_PSEPU
452 26 83.9 151 2 Q88L75_PSEPK
453 26 83.9 155 2 Q413G9_KINRA
454 26 83.9 156 2 Q4H2Y6_CIOIN
455 26 83.9 156 2 Q4C9K4_CROWT
456 26 83.9 156 2 Q4W540_CHITE
457 26 83.9 157 2 Q4DK86_9TRYR
458 26 83.9 160 2 Q4KHV0_PSEFS
459 26 83.9 169 2 Q3X413_9ACTN
460 26 83.9 174 1 YDSA_SCHPO
461 26 83.9 174 2 Q3CVD7_ALTAT
462 26 83.9 175 2 Q4F960_IPORA
463 26 83.9 178 2 Q3X435_9ACTN
464 26 83.9 179 2 Q5F6Y7_NEIGI
465 26 83.9 184 2 Q46PY1_RALEY
466 26 83.9 188 2 Q72G50_DESVH
467 26 83.9 189 2 Q4MTA2_BACCE
468 26 83.9 199 2 Q97B39_THEVO
469 26 83.9 200 2 Q5C2N9_SCHUA

```

```

Q4SV4 tetradon n
Q2T4P2 burkholderi
Q38B4 trypnosoma
Q9P2A8 homo sapien
Q4RMJ3 tetradon n
Q9U107 caenorhadi
Q8CFQ3 mus musculu
Q3ULE8 mus musculu
Q80TX8 mus musculu
Q60306 homo sapien
Q2SHZ2 habella che
Q93593 caenorhadi
Q40Y08 kinococcus
Q570B6 trypnosoma
Q4WKS1 aspergillus
Q5BGD9 aspergillus
Q4D8B9 trypnosoma
Q6P4S8 mus musculu
Q4DMW7 trypnosoma
Q4RXP0 tetradon n
Q7Q087 giardia lam
Q8R728 thermoaer
Q2VZM8 magnetospir
Q9KRZ6 vibrio chol
Q5BT57 schistosoma
Q58PS6 uncultured
P45030 saccharomyc
Q3AM77 synecococc
Q44SD5 chlorobium
Q6Z565 oryza sativ
Q9R6Q7 proteus mix
Q3X4U8 rubrobacter
Q9SFEZ arabidopsis
Q3RRUS ralsronia m
Q4N1J9 arthrobacte
Q2XVU1 bacterioph
Q3HLE6 bacillus an
Q2LIC3 bacillus an
Q2L1H6 bacillus an
Q3H1I0 bacillus an
Q218P3 bacillus an
Q3D1N5 streptococc
Q3DLF5 streptococc
Q548B4 streptococc
Q8DZ66 streptococc
Q8E4R2 streptococc
Q970Y7 streptococc
Q3VL18 pelodictyon
Q3ASB5 chlorobium
Q43GL7 chlorobium
Q3B6N1 pelodictyon
Q3AHZ9 chlorobium
Q8U1A4 virus plich
Q2XG16 pseudomonas
Q88L75 pseudomonas
Q413G9 kinococcus
Q4H2Y6 clona intes
Q4C9K4 crocospaer
Q4W540 chlorobium
Q4DK86 trypnosoma
Q4KHV0 pseudomonas
Q3X413 rubrobacter
Q4185 sechiosacch
Q3CVD7 pseudosalter
Q4F960 ipomeea bat
Q3X435 rubrobacter
Q5F6Y7 naiseeria g
Q46PY1 ralsronia g
Q72G50 desulfovibr
Q4MTA2 bacillus ce
Q97B39 thermoplasma
Q5C2N9 schistosoma

```

```

470 26 83.9 205 2 Q7NCK9_GLOVI
471 26 83.9 210 2 Q9H6X7_HUMAN
472 26 83.9 212 2 Q67WM8_ORYSA
473 26 83.9 212 2 Q89WM0_BRAJA
474 26 83.9 217 2 Q7FAT5_ORYSA
475 26 83.9 229 2 Q8W1I2_9CHLO
476 26 83.9 229 2 Q84FS6_9BACT
477 26 83.9 229 2 Q9L542_9BACT
478 26 83.9 230 2 Q3J287_RHOS4
479 26 83.9 230 2 Q9R699_9RHIZ
480 26 83.9 230 2 Q9R716_9RHIZ
481 26 83.9 231 2 Q9WME4_9RHIZ
482 26 83.9 231 2 Q8UEA7_9AGRT5
483 26 83.9 235 2 Q7WVZ1_PORCI
484 26 83.9 237 2 Q3WOK7_9ACTO
485 26 83.9 238 2 Q5AVB1_EMENT
486 26 83.9 238 2 Q3AMV7_SYNSC
487 26 83.9 240 2 Q9PCD9_9STRO
488 26 83.9 241 2 Q8Z6R4_9STRAM
489 26 83.9 246 2 Q4DC54_9TRYR
490 26 83.9 247 2 Q5KWE0_GEOXA
491 26 83.9 248 2 Q6WH44_BPXYA
492 26 83.9 250 2 Q8N1Z8_EMENT
493 26 83.9 250 2 Q5B3H5_EMENT
494 26 83.9 250 2 Q98PG7_MYCPU
495 26 83.9 252 1 PTGL_HUMAN
496 26 83.9 252 2 Q5SX19_MOUSE
497 26 83.9 255 2 Q8C1I4_MOUSE
498 26 83.9 259 2 Q4CNG1_9TRYR
499 26 83.9 259 2 Q3X1H3_9ACTN
500 26 83.9 261 2 Q17262_CABEL

```

ALIGNMENTS

```

RESULT 1
ID Q2SEI3_9GAMM PRELIMINARY; PRT; 324 AA.
AC Q2SEI3;
DC 24-JUN-2006, integrated into UniProtKB/TrEMBL.
DT 24-JUN-2006, sequence version 1.
DE Hypothetical protein.
DE Hypothetical protein.
GN ORFNames=HCH_04234;
OS Haehella chejuensis KCTC 2396.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Oceanospirillales;
OC Haehellaceae; Haehella.
OX NCBI_TaxID=349521;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=KCTC 2396;
RX PubMed=16352867; DOI=10.1093/nar/gki1016;
RA Jeong H., Yim J.H., Lee C., Choi S.-H., Park Y.K., Yoon S.H.,
RA Hur C.-G., Kang H.-Y., Kim D., Lee H.H., Park K.H., Park S.-H.,
RA Park H.-S., Lee H.K., Oh T.K., Kim J.F.;
RT "Genomic blueprint of Haehella chejuensis, a marine microbe producing
RT an alginate agent."
RL Nucleic Acids Res. 33:7066-7073(2005).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC
DR EMBL; CP000155; ABC30941.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ
SEQUENCE 324 AA; 36310 MW; E523A970CF1A081 CRC64;

```

```

Query Match 100.0%; Score 31; DB 2; Length 324;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 FSYRAR 6
|||||

```

Db 125 FSRYAR 130

RESULT 2

ID 0709K9 ANOGA PRELIMINARY; PRT; 338 AA.

AC 0709K9;

DT 15-DEC-2003, integrated into UniProtKB/TrEMBL.

DT 07-DEC-2004, sequence version 2.

DT 07-FEB-2006, entry version 7.

DE ENSANGP0000015636 (Fragment).

GN ORFNames=ENSANG00000013147;

OS Anopheles gambiae str. PEST.

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;

OC Anophelinae; Anopheles.

OX NCBI_TaxID=180454;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=PEST.

RG The Anopheles gambiae Sequence Committee;

RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.

CC -1- CAUTION: The sequence shown here is derived from an

CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is

CC preliminary data.

CC -----

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>

CC Distributed under the Creative Commons Attribution-NoDerivs License

CC -----

DR EMBL: AAB0106900; EAA09433.2; -; Genomic_DNA.

DR NON_TER 1

DR SEQUENCE 338 AA; 39557 MW; B6534CB28EFA3737 CRC64;

Query Match 100.0%; Score 31; DB 2; Length 338;

Best Local Similarity 100.0%; Pred. No. 86;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRYAR 6

Db 277 FSRYAR 282

RESULT 3

PAAE_ECOLI

ID PAAE_ECOLI STANDARD; PRT; 356 AA.

AC P76081; O53013; P77233;

DT 27-APR-2001, integrated into UniProtKB/Swiss-Prot.

DT 01-FEB-1997, sequence version 1.

DT 07-MAR-2006, entry version 56.

DE Probable phenylacetic acid degradation NADH oxidoreductase paab

DE (EC 1.-.-.-).

GN Name=paab; Ordered locus names=b1392;

OS Escherichia coli.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Escherichia.

OX NCBI_TaxID=562;

RN [1]

RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].

RC STRAIN=W / ATCC 11105;

RX MEDLINE=98421522; PubMed=9748275; DOI=10.1074/jbc.273.40.25974;

RA Ferrandez A., Minambres B., Garcia B., Olvera E.R., Luengo J.M.,

RA Garcia J.L., Diaz E.;

RT "Catabolism of phenylacetic acid in Escherichia coli. Characterization

RT of a new aerobic hybrid pathway.";

RL J. Biol. Chem. 273:25974-25986(1998).

RN [2]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=K12 / MG1655;

RX MEDLINE=97426617; PubMed=9278503; DOI=10.1126/science.277.5331.1453;

RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,

RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,

RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,

RA Mau B., Shao Y.;

RT "The complete genome sequence of Escherichia coli K-12.";

RL Science 277:1453-1474(1997).

RN [3]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=K12;

RX MEDLINE=97251357; PubMed=9097039; DOI=10.1093/dnares/3.6.363;

RA Alpha H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,

RA Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,

RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,

RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T., Saito N.,

RA Sempel G., Seki Y., Sivasubram S., Tagami H., Takeda J.,

RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horituchi T.;

RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome

RT corresponding to the 28.0-40.1 min region on the linkage map.";

RL DNA Res. 3:363-377(1996).

RN [4]

RP TRANSCRIPTIONAL REGULATION.

RX MEDLINE=20229831; PubMed=10766858; DOI=10.1074/jbc.275.16.12214;

RA Ferrandez A., Garcia J.L., Diaz E.;

RT "Transcriptional regulation of the divergent paa catabolic operons for

RT phenylacetic acid degradation in Escherichia coli.";

RL J. Biol. Chem. 275:12214-12222(2000).

CC -1- FUNCTION: May be part of a multicomponent oxygenase involved in

CC phenylacetyl-CoA hydroxylation.

CC -1- COFACTOR: Binds 1 2Fe-2S cluster (By similarity).

CC -1- PATHWAY: Phenylacetic acid aerobic catabolism.

CC -1- INDUCTION: Activated by cAMP receptor protein (CRP) and

CC integration host factor (IHF). Inhibited by paaX.

CC -1- SIMILARITY: In the N-terminal section, belongs to the FAD-binding

CC oxidoreductase type 6 family.

CC -1- SIMILARITY: Contains 1 2Fe-2S ferredoxin-type domain.

CC -----

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>

CC Distributed under the Creative Commons Attribution-NoDerivs License

CC -----

DR EMBL: X97452; CA66094.1; -; Genomic_DNA.

DR EMBL: U00096; AAC74474.1; -; Genomic_DNA.

DR EMBL: D90777; BAA14998.1; ALT_INIT; Genomic_DNA.

DR EMBL: D90778; BAA15003.1; ALT_INIT; Genomic_DNA.

DR PIR: C64890; C64890.

DR HSSP: P06543; 1J7B.

DR GenomeReviews: U00096_GR; b1392.

DR ECHOBASE: EB3502; -.

DR EcoGene: EGI3739; paaE.

DR BioCyc: EcoCyc:G6713-MONOMER; -.

DR LinkHub: P76081; -.

DR InterPro: IPR006058; 2Fe2S fd BS.

DR InterPro: IPR01041; Ferredoxin.

DR InterPro: IPR012675; Ferredoxin fold.

DR InterPro: IPR001709; FPN_cyt_redctse.

DR InterPro: IPR008333; Oxred_FAD_bd.

DR InterPro: IPR001433; Oxred_FAD_NAD_bd.

DR InterPro: IPR011884; PA CoA Oase5.

DR InterPro: IPR001221; Phe hydroxylase.

DR Pfam: PF00970; FAD_binding_6; 1.

DR Pfam: PF00111; Fe2; 1.

DR Pfam: PF00175; NAD_binding_1; 1.

DR PRINTS: PR00371; FENCR.

DR PRINTS: PR00410; PHEHYDLASE.

DR TIGRFAMs: TIGR02160; PA CoA Oxy5; 1.

DR PROSITE: PS1085; 2FE2S_FER_2; 1.

DR PROSITE: PS00197; 2FE2S_FER_2; 1.

DR 2Fe-2S; Complete proteome; Electron transport; FAD; Flavoprotein;

KW Iron; Iron-sulfur; Metal-binding; NAD; Oxidoreductase; Transport.

FT CHAIN 1 356

FT NADH oxidoreductase paab.

```

FT DOMAIN 262 354 /FTId=PRO.0000058163.
FT REGION 112 228 2Fe-2S ferredoxin-type.
FT METAL 299 299 Oxidoreductase (Potential).
FT METAL 304 304 Iron-sulfur (2Fe-2S) (By similarity).
FT METAL 307 307 Iron-sulfur (2Fe-2S) (By similarity).
FT METAL 337 337 Iron-sulfur (2Fe-2S) (By similarity).
FT VARIANT 14 14 S -> P (in strain: W).
FT VARIANT 169 169 P -> S (in strain: W).
FT VARIANT 224 227 DAET -> ETEA (in strain: W).
SQ SEQUENCE 356 AA; 39320 MW; D719C1CA81DA5FFA CRC64;

Query Match 100.0%; Score 31; DB 1; Length 356;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSRYAR 6
Db 80 FSRYAR 85

RESULT 4
MURG_CHLN STANDARD; PRT; 357 AA.
ID MURG_CHLN
AC Q92702; Q9J027;
DT 30-MAY-2000, integrated into UniProtKB/Swiss-Prot.
DT 01-MAY-1999, sequence version 1.
DT 07-MAR-2006, entry version 36.
DE UDP-N-acetylglucosamine-N-acetylmuramyl-(pentapeptide)
DE pyrophosphoryl-undecaprenol N-acetylglucosamine transferase
DE (EC 2.4.1.22) (Undecaprenyl-PP-MurNAc-pentapeptide-UDPglcNAc
DE transferase).
GN Name=murg; OrderedLocustNames=CPN0904, CP0962, CPB0936;
OS Chlamydia pneumoniae (Chlamydiophila pneumoniae).
OX Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydiophila.
NCBI_TaxID=83558;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=J138;
RA MEDLINE=9920606; PubMed=10192388; DOI=10.1038/7716;
RA Kaiman S., Mitchell W.P., Marathe R., Lammel C.J., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.,
RA "Comparative genomes of Chlamydia pneumoniae and C. trachomatis";
RL Nat. Genet. 21:385-389(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=AR39;
RA MEDLINE=20150255; PubMed=10684935; DOI=10.1093/nar/28.6.1397;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heideberg J.F.,
RA White O., Hickey E.K., Peterson J.D., Utecherback T.R., Berry K.J.,
RA Baas S., Linkey K.D., Weidman J.F., Khouri H.M., Craven B., Bowman C.,
RA Dodson R.J., Ginn W.L., Nelson W.C., DeBoy R.T., Kolonay J.F.,
RA McLacty G., Salzberg S.L., Eisen J.A., Fraser C.M.;
RA "Genome sequences of Chlamydia trachomatis Moyn and Chlamydia
RA pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=J138;
RA MEDLINE=20330349; PubMed=10871362; DOI=10.1093/nar/28.12.2311;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RA "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RA from Japan and CWL029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=TW-183;
RA Geng M.M., Schumacher A., Muehlhofer I., Bensch K.W., Schaefer K.P.,
RA Schneider S., Pohl T., Essig A., Marre R., Melchers K.;
RA "The genome sequence of Chlamydia pneumoniae TW18 and comparison with
RA other Chlamydia strains based on whole genome sequence analysis.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

```

```

CC -1- FUNCTION: Cell wall formation. Catalyzes the transfer of a GlcNAc
CC subunit on undecaprenyl-pyrophosphoryl-MurNAc-pentapeptide (lipid
CC intermediate II) to form undecaprenyl-pyrophosphoryl-MurNAc-
CC (pentapeptide)GlcNAc (lipid intermediate II) (By similarity).
CC -1- CATALYTIC ACTIVITY: UDP-N-acetylglucosamine + Mur2Ac(oyl-D-Ala-
CC gamma-D-Glu-L-Lys-D-Ala-D-Ala)-diphosphoundecaprenol = UDP +
CC GlcNAc-(1->4)-Mur2Ac(oyl-L-Ala-gamma-D-Glu-L-Lys-D-Ala-D-Ala)-
CC diphosphoundecaprenol.
CC -1- PATHWAY: Peptidoglycan biosynthesis; last step.
CC -1- SUBCELLULAR LOCATION: Bacterial cell inner membrane; peripheral
CC membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the glycosyltransferase 28 family. MurG
CC subfamily.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL, AE001670; MAD19042.1; -; Genomic DNA.
DR EMBL, AE002254; AAF38742.1; -; Genomic DNA.
DR EMBL, BA000008; BAA99112.1; -; Genomic DNA.
DR EMBL, AE017160; AAP98865.1; ALT_INIT; Genomic DNA.
DR PIR, C72022; C72022.
DR PIR, F86603; F86603.
DR GenomeReviews; AE002161.GR; CP0962.
DR GenomeReviews; AE009440.GR; CPB0936.
DR GenomeReviews; AE001363.GR; CPN0904.
DR GenomeReviews; BA000008.GR; murg.
DR TIGR, CP0962; -.
DR BioCyc; CPN0115711.CPN0962-MONOMER; -.
DR BioCyc; CPN0115713.CPN0904-MONOMER; -.
DR BioCyc; CPN0138677.CPN0904-MONOMER; -.
DR BioCyc; CPN0182082.CPB0936-MONOMER; -.
DR HAMAP; MF 00033; -; 1.
DR InterPro; IPR007235; Glyco tran_28_C.
DR InterPro; IPR004276; Glyco_trans_26.
DR InterPro; IPR006009; Murg.
DR Pfam; PF04101; Glyco_tran_28_C; 1.
DR Pfam; PF03033; Glyco_transf_28; 1.
DR TIGRFAMs; TIGR01133; murg; 1.
KW Cell cycle; Cell division; Cell shape; Cell wall; Complete proteome;
KW Glycosyltransferase; Inner membrane; Membrane;
KW Peptidoglycan synthesis; Transferase.
FT CHAIN 1 357
FT UDP-N-acetylglucosamine--N-acetylmuramyl-
FT (pentapeptide) pyrophosphoryl-
FT undecaprenol N-acetylglucosamine
FT transferase.
FT /FTId=PRO.0000109161.
FT SQ SEQUENCE 357 AA; 39420 MW; DEB8350CF168BC42 CRC64;

Query Match 100.0%; Score 31; DB 1; Length 357;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSRYAR 6
Db 135 FSRYAR 140

RESULT 5
Q5SMX2_ORYSA PRELIMINARY; PRT; 412 AA.
ID Q5SMX2_ORYSA
AC Q5SMX2;
DT 21-DEC-2004, integrated into UniProtKB/TrEMBL.
DT 21-DEC-2004, sequence version 1.
DT 07-FEB-2006, entry version 7.
DE Basic helix-loop-helix protein-like.
GN Name=P04980; P04980.25;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BPF clade;
OC Ehrhartoideae; Oryzae; Oryza.
NCBI_TaxID=39947;
RN [1]

```



```

RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=2237376; Pubmed=12447438; DOI=10.1038/nature01184;
RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
RA Wu J., Nishimura Y., Cheng Z., Nagamura Y., Antonio B.A., Katanamori H.,
RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi C.,
RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
RA Hijikita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,
RA Ikono M., Ito S., Ito T., Ito Y., Iwabuchi A., Kamiya K.,
RA Karaawa M., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
RA Machita K., Maehara T., Mizuno H., Mizuyaashi T., Mukai Y.,
RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai K., Shibata M.,
RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,
RA Maki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,
RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.-I., Eun M.-Y.,
RA Yano M., Jiang J., Gyojohori T.;
RT "The genome sequence and structure of rice chromosome 1."
RL Nature 420:312-316(2002).
CC -1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL; AP003417; BAD72431.1; -; Genomic_DNA.
CC
DR Gramene; OS5MX2; -.
CC
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0030528; F:transcription regulator activity; IEA.
DR GO; GO:0045449; P:regulation of transcription; IEA.
DR InterPro; IPR001092; HLH_basic.
DR InterPro; IPR015598; HLH_DNA_bd.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS50868; HLH; 1.
DR SQ SEQUENCE 412 AA; 43709 MW; 644AA46084EB40B2 CRC64;
QY 1 FSRYPAR 6
Db 117 FSRYPAR 122
Query Match 100.0%; Score 31; DB 2; Length 412;
Best Local Similarity 100.0%; Pred. No. 1.le+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

DR EMBL; BA000012; BAB48618.1; -; Genomic_DNA.
DR BioCyc; MLOT381; ML1183-MONOMER; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR InterPro; IPR011761; ATP_GASP.
DR PROSITE; PS50975; ATP_GASP; 1.
KM ATP-binding; Complete proteome; Nucleotide-binding.
SQ SEQUENCE 427 AA; 46722 MW; F4435AC1EDC7F645 CRC64;
QY 1 FSRYPAR 6
Db 53 FSRYPAR 58
Query Match 100.0%; Score 31; DB 2; Length 427;
Best Local Similarity 100.0%; Pred. No. 1.le+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 FSRYAR 6
    |||||
Db 147 FSRYAR 152

RESULT 8
Q62CX5 BURMA PRELIMINARY; PRT; 447 AA.
ID 062CX5 BURMA PRELIMINARY; PRT; 447 AA.
AC 062CX5;
DT 25-OCT-2004, integrated into UniProtKB/TrEMBL.
DT 25-OCT-2004, sequence version 1.
DT 07-FEB-2006, entry version 16.
DE Amino acid permealase.
GN OrderedLocustNames=BMAA0717;
OS Burkholderia mallei (Pseudomonas mallei).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=13373;
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP STRAIN=ATCC 23344;
RX PubMed=15377793; DOI=10.1073/pnas.0403306101;
RA Niernan W.C., Deshazer D., Kim H.S., Tectelin H., Nelson K.E.,
RA Feldblum T.V., Ulrich R.L., Romning C.M., Birkac L.M.,
RA Daugherty S.C., Davidsen T.D., DeBoy R.T., Dimitrov G., Dodson R.J.,
RA Durkin A.S., Gwinn M.L., Haft D.H., Khouri H.M., Kolonay J.F.,
RA Madupu R., Mohammed Y., Nelson W.C., Radune D., Romero C.M.,
RA Sarría S., Selengut J., Shambin C., Sullivan S.A., White O., Yu Y.,
RA Zafar N., Zhou L., Fraser C.M.;
RT "Structural flexibility in the Burkholderia mallei genome.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14246-14251(2004).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL; CP000011; AAU46888.1; -; Genomic DNA.
DR TIGR; BMAA0717; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005279; F:amino acid-polyamine transporter activity; IEA.
DR GO; GO:0006865; P:amino acid transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR002293; AA/rel_permease1.
DR InterPro; IPR004840; AAC permease.
DR InterPro; IPR004841; Permease region.
DR PANTHER; PTHR11785; AA/rel_permease1; 1.
DR Pfam; PF00324; AA_permease; 1.
DR PROSITE; PS00218; AMINO_ACID_PERMEASE_1; 1.
DR Complete proteome.
SQ SEQUENCE 447 AA; 48683 MW; A209ED83F6C1D427 CRC64;

Query Match 100.0%; Score 31; DB 2; Length 447;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 FSRYAR 6
    |||||
Db 76 FSRYAR 81

RESULT 9
Q45AY3 BURK PRELIMINARY; PRT; 460 AA.
ID Q45AY3 BURK PRELIMINARY; PRT; 460 AA.
AC Q45AY3;
DT 13-SEP-2005, integrated into UniProtKB/TrEMBL.
DT 13-SEP-2005, sequence version 1.
DT 07-FEB-2006, entry version 5.
DE Amino acid permealase-associated region.
GN ORFNames=BcendraFT_5191;
OS Burkholderia cenocepacia AU 1054.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.

```

```

OC Burkholderiaceae; Burkholderia; Burkholderia cenocepacia complex.
OX NCBI_TaxID=331271;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=AU 1054;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RA Hammon N., Israni S., Pfluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Burkholderia cenocepacia
RT AU 1054.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=AU 1054;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Lamer F., Land M.;
RT "Annotation of the draft genome assembly of Burkholderia cenocepacia
RT AU 1054." (JUN-2005) to the EMBL/GenBank/DBJ databases.
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL; AAH10100004; EAM13792.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005279; F:amino acid-polyamine transporter activity; IEA.
DR GO; GO:0006865; P:amino acid transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR002293; AA/rel_permease1.
DR InterPro; IPR004840; AAC permease.
DR Pfam; PF00324; AA_permease; 1.
DR PROSITE; PS00218; AMINO_ACID_PERMEASE_1; 1.
DR Complete proteome.
SQ SEQUENCE 460 AA; 49648 MW; 68760B595F03369D CRC64;

Query Match 100.0%; Score 31; DB 2; Length 460;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 FSRYAR 6
    |||||
Db 76 FSRYAR 81

RESULT 10
Q4LTZ2 BURK PRELIMINARY; PRT; 460 AA.
ID Q4LTZ2 BURK PRELIMINARY; PRT; 460 AA.
AC Q4LTZ2;
DT 02-AUG-2005, integrated into UniProtKB/TrEMBL.
DT 02-AUG-2005, sequence version 1.
DT 07-FEB-2006, entry version 5.
DE Amino acid permealase-associated region.
GN ORFNames=Bcen2424DRAFT_4131;
OS Burkholderia cenocepacia H12424.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia; Burkholderia cenocepacia complex.
OX NCBI_TaxID=331272;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=H12424;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RA Hammon N., Israni S., Pfluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Burkholderia cenocepacia
RT H12424.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=H12424;
RG US DOE Joint Genome Institute (JGI-ORNL);

```

```

RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Burkholderia cenocepacia
RT HI2424."
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL: AAHL0100016; FAM19436.1; -; Genomic DNA.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0005279; F:amino acid-polyamine transporter activity; IEA.
DR GO: GO:0006865; P:amino acid transport; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR002293; AA/rel_permease1.
DR InterPro: IPR004840; AAC_permease.
DR PANTHER: PTHR11785; AA/rel_permease1; 1.
DR Pfam: PF00324; AA_permease; 1.
DR PROSITE: PS00218; AMINO_ACID_PERMEASE_1; 1.
SQ SEQUENCE 460 AA; 49648 MW; 68780B9B9F03369D CRC64;

Query Match 100.0%; Score 31; DB 2; Length 460;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSRYAR 6
Db 76 FSRYAR 81

RESULT 11
Q395V1 BUR53 PRELIMINARY; PRT; 460 AA.
AC Q395V1
DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.
DT 22-NOV-2005, sequence version 1.
DT 21-FEB-2006, entry version 5.
DE Amino acid transporter.
OS Burkholderia sp. (strain 383) (Burkholderia cepacia (strain ATCC 17760
/ NCIB 9086 / R18194)).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia; Burkholderia cepacia complex.
OX NCBI_TaxID=269483;
[1]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RG US DOE Joint Genome Institute.
RA Copeland A., Lucas S., Lapidus A., Barry K., Deter J.C., Glavina T.,
RA Hammon N., Israel S., Pittluck S., Chain P., Malfatti S., Shin M.,
RA Verger L., Schmutz J., Larimer F., Land M., Kyrides N., Lykidis A.,
RA Richardson P.;
RT "Complete sequence of chromosome 2 of Burkholderia sp. 383."
RL Submitted (OCT-2005) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL: CP000152; ABH11760.1; -; Genomic DNA.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0005279; F:amino acid-polyamine transporter activity; IEA.
DR GO: GO:0006865; P:amino acid transport; IEA.
DR GO: GO:0006810; P:transport; IEA.
KW Complete proteome.
SQ SEQUENCE 460 AA; 49728 MW; DGBBA9AF2CD5155 CRC64;

Query Match 100.0%; Score 31; DB 2; Length 460;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 FSRYAR 6
Db 76 FSRYAR 81

RESULT 12
Q2T424 BURTH PRELIMINARY; PRT; 461 AA.
ID Q2T424 BURTH
AC Q2T424
DT 24-JAN-2006, integrated into UniProtKB/TrEMBL.
DT 24-JAN-2006, sequence version 1.
DT 07-FEB-2006, entry version 2.
DE Amino acid permease.
GN ORFNames=BTB_111560;
OS Burkholderia thailandensis B264.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia; pseudomallei group.
OX NCBI_TaxID=271848;
[1]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=E264.
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Palmer N., Dodson R.,
RA Hickey B.K., Gwinn M., Dougherty B., Fleischmann R.D., Richardson D.,
RA Peterson J., Kervaege A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van-Vugt R., Adams M.D., Gocayne J.D., Weidman C., Ullrich T.,
RA Wathey L., McDonald L., Artiach P., Bowman C., Garland S., Fujii C.,
RA Cotton M.D., Horst K., Tomb J.-F., Roberts K., Hatch B., Smith H.O.,
RA Venter J.C.;
RL Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL: CP000085; ABC5479.1; -; Genomic DNA.
DR GO: GO:0005279; F:amino acid-polyamine transporter activity; IEA.
SQ SEQUENCE 461 AA; 49957 MW; 7F1195E772F28B65 CRC64;

Query Match 100.0%; Score 31; DB 2; Length 461;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSRYAR 6
Db 76 FSRYAR 81

RESULT 13
Q3JFR3 BURP1 PRELIMINARY; PRT; 461 AA.
ID Q3JFR3 BURP1
AC Q3JFR3
DT 08-NOV-2005, integrated into UniProtKB/TrEMBL.
DT 08-NOV-2005, sequence version 1.
DT 21-FEB-2006, entry version 5.
DE Amino acid permease.
GN Name=proX; OrderedlocusNames=BURP1_1710b_A2440;
OS Burkholderia pseudomallei (strain 1710b)
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia; pseudomallei group.
OX NCBI_TaxID=320372;
[1]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP Moore D.E., Nieman W.C.;
RL Submitted (SEP-2005) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL: CP000125; ABA52590.1; -; Genomic DNA.
DR TIGR: BURP1_1710b_A2440; -;
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0005279; F:amino acid-polyamine transporter activity; IEA.
DR GO: GO:0006865; P:amino acid transport; IEA.

```

DR GO: GO:0006810; P:transport; IEA.
 DR InterPro: IPR002293; AA:rel_permease1.
 DR InterPro: IPR004840; AAC_permease.
 DR InterPro: IPR004841; Permease_region.
 DR Pfam: PF00324; AA_permease_1; 1.
 DR PROSITE: PS00218; AMINO_ACID_PERMEASE_1; 1.
 DR Complete proteome.
 SQ SEQUENCE 461 AA; 50051 MW; 2DBA117A4942B62E CRC64;
 Query Match 100.0%; Score 31; DB 2; Length 461;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 FSRYAR 6
 |||||
 76 FSRYAR 61
 RESULT 14
 O63M15 BURPS PRELIMINARY; PRT; 461 AA.
 ID O63M15 BURPS
 AC O63M15;
 DT 25-OCT-2004, integrated into UniProtKB/TrEMBL.
 DT 25-OCT-2004, sequence version 1.
 DT 07-FEB-2006, entry version 17.
 DE Putative proline-specific permease.
 GN Name=ProY; OrderedLocustNames=BPS0845;
 OS Burkholderia pseudomallei (Pseudomonas pseudomallei).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Burkholderiaceae; Burkholderia; pseudomallei group.
 NCBI_TaxID=28450;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=K96243;
 RX PubMed=15377794; DOI=10.1073/pnas.0403302101;
 RA Holden M.T.G., Fittell R.W., Peacock S.O., Cerdano-Tarraga A.-M.,
 Aktins T., Crossman L.C., Pitt T., Churcher C., Mungall K.L.,
 Bentley S.D., Sabathia M., Thomson N.R., Bason N., Beacham I.R.,
 Brooks K., Brown K.A., Brown N.F., Challis G.L., Cherevach I.,
 Chillingworth T., Cronin A., Crosser B., Davis P., Desnazer D.,
 Fellwell T., Fraser A., Hance Z., Hauser H., Holtroyd S., Jagels K.,
 Keith K.E., Maddison M., Moule S., Price C., Quail M.A.,
 Rabinowitsch E., Rutherford K., Sanders M., Simmonds M.,
 Songswilvit S., Stevens K., Tumapa S., Vesaratchavee M.,
 Whitehead S., Yeates C., Barrett B.G., Oyston P.C.F., Parkhill J.,
 "Genomic plasticity of the causative agent of melioidosis,
 Burkholderia pseudomallei";
 RL Proc. Natl. Acad. Sci. U.S.A. 101:14240-14245(2004).
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs license
 CC -----
 DR EMBL: BX571966; CAH8307.1; -; Genomic_DNA.
 DR GO: GO:0016021; C:Integral to membrane; IEA.
 DR GO: GO:0016020; C:membrane; IEA.
 DR GO: GO:0005279; F:amino acid-polyamine transporter activity; IEA.
 DR GO: GO:0006865; P:amino acid transport; IEA.
 DR GO: GO:0006810; P:transport; IEA.
 DR InterPro: IPR002293; AA:rel_permease1.
 DR InterPro: IPR004840; AAC_permease.
 DR InterPro: IPR004841; Permease_region.
 DR PANTHER: PTHR11785; AA:rel_permease1; 1.
 DR Pfam: PF00324; AA_permease1; 1.
 DR PROSITE: PS00218; AMINO_ACID_PERMEASE_1; 1.
 DR Complete proteome.
 SQ SEQUENCE 461 AA; 50021 MW; D7BAE75B033B15F CRC64;
 Query Match 100.0%; Score 31; DB 2; Length 461;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 FSRYAR 6
 |||||

DB 76 FSRYAR 61
 RESULT 15
 O34618 BACSU PRELIMINARY; PRT; 463 AA.
 ID O34618 BACSU
 AC O34618; Q795P5;
 DT 01-JAN-1998, integrated into UniProtKB/TrEMBL.
 DT 01-JAN-1998, sequence version 1.
 DT 07-FEB-2006, entry version 30.
 DE Amino acid transporter (YtnA protein).
 GN Name=YtnA; OrderedLocustNames=BSU030530;
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 NCBI_TaxID=1423;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=98048467; PubMed=9387221;
 RA Lapidus A., Galleron N., Sorokin A., Ehrlich S.-D.;
 RT "Sequencing and functional annotation of the Bacillus subtilis genes
 in the 200 kb rnb-dnaB region";
 RL Microbiology 143:3431-3441(1997).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Lapidus A., Galleron N., Sorokin A., Ehrlich S.-D.;
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377; DOI=10.1038/36786;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 Azevedo V., Bertaino M.G., Bessieres P., Bolotin A., Borchert S.,
 Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
 Choi S.-K., Codani J.-J., Connerton I.F., Cummings N.J., Daniel R.A.,
 Demizot F., Devine K.M., Duesterhoeft A., Ehrlich S.D., Emerson P.T.,
 Entian K.-D., Errington J., Fabret C., Ferrati E., Fougier D.,
 Fritz C., Fujita M., Fujita Y., Fuma S., Gallizzi A., Galleron N.,
 Gilm S.-Y., Glaeser P., Goffeau A., Golightly E.J., Grandi G.,
 Giuseppe G., Guy B.J., Haga K., Haelech J., Harwood C.R., Henaut A.,
 Hilbert H., Holsappel S., Hosono S., Hulio M.-F., Ilaya M.,
 Jones L.W., Joris B., Karamata D., Kasaara Y., Kleier-Blanchard M.,
 Klein C., Kobayashi Y., Koetter P., Koningsstein G., Krogh S.,
 Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J.,
 Lazarevic V., Lee S.-M., Levine A., Liu H., Masuda S., Mauel C.,
 Medigue C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S.,
 Noback M., Noone D., O'Reilly M., Ogawa K., Ogihara A., Oudega B.,
 Park S.-H., Parro V., Pohl T.M., Portetelle D., Porwollik S.,
 Prescott A.M., Prescan E., Pujic P., Purnelle B., Rapoport G.,
 Ray M., Reynolds S., Rieger M., Rivolta C., Rocha E., Roche B.,
 Rose M., Sadleir Y., Sato T., Scanlan E., Schleich S., Schroeter R.,
 Scoffone F., Sekiguchi J., Sekowska A., Serr S.J., Serron P.,
 Shin B.-S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H.,
 Takemaru K., Takeuchi M., Tamakoshi A., Tanaka T., Terpestra P.,
 Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vannier F.,
 Vassarotti A., Viari A., Wambut R., Wedler E., Wedler H.,
 Weitzneger T., Winters P., Wipat A., Yamamoto H., Yamane K.,
 Yasunoto K., Yata K., Yoshida K., Yoshikawa H.-F., Zumstein E.,
 Yoshikawa H., Zancin A.;
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 subtilis";
 RL Nature 390:249-256(1997).
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs license
 CC -----
 DR EMBL: AF008220; AAC00244.1; -; Genomic_DNA.
 DR EMBL: Z99119; CAB15031.1; -; Genomic_DNA.
 DR PIR: C69997; C69997.
 DR BioCyc: BSU0423;BSU048-MONOMER; -;
 DR GO: GO:0016021; C:Integral to membrane; IEA.
 DR GO: GO:0016020; C:membrane; IEA.
 DR GO: GO:0005279; F:amino acid-polyamine transporter activity; IEA.

DR GO: GO:0006865; P:amino acid transport; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR002293; AA/rel_permease1.
DR InterPro: IPR004840; AAC_permease.
DR InterPro: IPR004841; Permease_region.
DR PANTHER: PTHR11785; AA/rel_permease1; 1.
DR Pfam: PF00324; AA_permease; 1.
DR PROSITE: PS00218; AMINO_ACID_PERMEASE_1; UNKNOWN_1.
DR Complete proteome; Transmembrane; transport.
KW SEQUENCE 463 AA; 50328 MW; B21F9F40AB1C7180 CRC64;
SQ
Query Match 100.0%; Score 31; DB 2; Length 463;
Best Local Similarity 100.0%; Pred. NO. 1.2e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 FSRYAR 6
Db 76 FSRYAR 81
RESULT 16
Q2STX5 BURTH PRELIMINARY; PRT; 468 AA.
ID Q2STX5 BURTH PRELIMINARY; PRT; 468 AA.
AC Q2STX5;
DT 24-JAN-2006, integrated into UniProtKB/TrEMBL.
DT 24-JAN-2006, sequence version 1.
DT 07-FEB-2006, entry version 2.
DE Amino acid permease.
GN ORFNames=BTH_13129;
OS Burkholderia thailandensis E264.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia; pseudomallei group.
OX NCBI_TaxID=271848;
RN [1]
NP NUCLEOTIDE SEQUENCE.
RC STRAIN=E264;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Palmer N., Dodson R.,
RA Hickey E.K., Gwin M., Dougherty B., Fleischmann R.D., Richardson D.,
RA Peterson J., Kervilange A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van-Vugt R., Adams M.D., Gocayne J.D., Weidman J., Utechtack T.,
RA Matthew L., McDonald L., Artach P., Bowman C., Garland S., Fujii C.,
RA Cotton M.D., Horst K., Tomb J.-F., Roberts K., Hatch B., Smith H.O.,
RA Venter J.C.;
RL Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL: CP000086; ABC38752.1; -; Genomic DNA.
DR SEQUENCE 468 AA; 50779 MW; A905AA5CB57B7284 CRC64;
SQ
Query Match 100.0%; Score 31; DB 2; Length 468;
Best Local Similarity 100.0%; Pred. NO. 1.2e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 FSRYAR 6
Db 78 FSRYAR 83
RESULT 17
Q3JY99 BURP1 PRELIMINARY; PRT; 468 AA.
ID Q3JY99 BURP1 PRELIMINARY; PRT; 468 AA.
AC Q3JY99;
DT 08-NOV-2005, integrated into UniProtKB/TrEMBL.
DT 08-NOV-2005, sequence version 1.
DT 21-FEB-2006, entry version 5.
DE Amino acid permease.
GN OrderedListNames=BURPS1710b_0036;
OS Burkholderia pseudomallei (strain 1710b).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia; pseudomallei group.

OX NCBI_TaxID=320372;
RN [1]
NP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RA Woods D.E., Nielsen W.C.;
RL Submitted (SEP-2005) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL: CP000124; ABA50182.1; -; Genomic DNA.
DR TIGR: BURPS1710b_0036; -;
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0005279; F:amino acid-polyamine transporter activity; IEA.
DR GO: GO:0006865; P:amino acid transport; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR002293; AA/rel_permease1.
DR InterPro: IPR004840; AAC_permease.
DR InterPro: IPR004841; Permease_region.
DR Pfam: PF00324; AA_permease; 1.
DR PROSITE: PS00218; AMINO_ACID_PERMEASE_1; 1.
DR Complete proteome.
KW SEQUENCE 468 AA; 50889 MW; EBEFA2C3CE681110 CRC64;
SQ
Query Match 100.0%; Score 31; DB 2; Length 468;
Best Local Similarity 100.0%; Pred. NO. 1.2e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 FSRYAR 6
Db 78 FSRYAR 83
RESULT 18
Q62G62 BURMA
ID Q62G62 BURMA PRELIMINARY; PRT; 468 AA.
AC Q62G62;
DT 25-OCT-2004, integrated into UniProtKB/TrEMBL.
DT 25-OCT-2004, sequence version 1.
DT 07-FEB-2006, entry version 16.
DE Amino acid permease.
GN OrderedListNames=BMA2796;
OS Burkholderia mallei (Pseudomonas mallei).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=13373;
RN [1]
NP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=ATCC 23344;
RX PubMed=15377793; DOI=10.1073/pnas.040306101;
RA Nierman W.C., Deshazer D., Kim H.S., Tettelin H., Nelson K.E.,
RA Feldblyum T.V., Ulrich R.L., Rinning C.M., Brinkac L.M.,
RA Dougherty S.C., Davidson T.D., DeBoy R.T., Dimitrov G., Dodson R.J.,
RA Durkin A.S., Gwin M.L., Haft D.H., Khouri H.M., Kolonay J.F.,
RA Madupu R., Mohammed Y., Nelson W.C., Radune D., Romero C.M.,
RA Sarría S., Selengut J., Shamlin C., Sullivan S.A., White O., Yu Y.,
RA Zafar N., Zhou L., Fraser C.M.;
RL "Structural flexibility in the Burkholderia mallei genome";
RT Proc. Natl. Acad. Sci. U.S.A. 101:14246-14251(2004).
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL: CP000010; AAU47957.1; -; Genomic DNA.
DR TIGR: BMA2796; -;
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0005279; F:amino acid-polyamine transporter activity; IEA.
DR GO: GO:0006865; P:amino acid transport; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR002293; AA/rel_permease1.
DR InterPro: IPR004840; AAC_permease.
DR InterPro: IPR004841; Permease_region.

DR PANTHER; PTHR11785; AA/rel_permease1; 1.
 DR Pfam; PF00324; AA_permease; 1.
 DR PROSITE; PS00218; AMINO_ACID_PERMEASE_1; 1.
 KM Complete proteome.
 SQ SEQUENCE 468 AA; 50949 MW; AAEB96964CBBBCB CRC64;

Query Match 100.0%; Score 31; DB 2; Length 468;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRYAR 6
 |||||
 Db 78 FSRYAR 83

RESULT 19
 ID O63PW6_BURPS PRELIMINARY; PRT; 468 AA.
 AC O63PW6;
 DT 25-OCT-2004, integrated into UniProtKB/TrEMBL.
 DT 25-OCT-2004, sequence version 1.
 DT 07-FEB-2006, entry version 16.
 DE Putative amino acid permease.
 GN OrderedLocustNames=BSJ3255;
 OS Burkholderia pseudomallei (Pseudomonas pseudomallei).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Burkholderiaceae; Burkholderia; pseudomallei group.
 OX NCBI_TaxId=28450;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=K96243;
 RX PubMed=15377794; DOI=10.1073/pnas.0403021011;
 RA Holden M.T.G., Tilball R.W., Peacock S.J., Cerdano-Tarraga A.-M.,
 Ahtinen T., Crossman L.C., Pitt T., Churcher C., Mungall K.L.,
 Bentley S.D., Sebahia M., Thomson N.R., Bason N., Beacham I.R.,
 Brooks K., Brown K.A., Brown N.F., Challis G.L., Chevreton I.,
 Chillingworth T., Cronin A., Crosser B., Davis P., Desnazer D.,
 Fellwell T., Fraser A., Hance Z., Hauser H., Holtroyd S., Jagels K.,
 Keith K.E., Maddison M., Moule S., Price C., Quail M.A.,
 Rabinowitsch E., Rutherford K., Sanders M., Simmonds M.,
 Songshivalai S., Stevens K., Tumapa S., Vesaratchavee M.,
 Whitehead S., Yeats C., Barrett B.G., Oyston P.C.F., Parkhill J.;
 RT "Genomic plasticity of the causative agent of melioidosis,
 Burkholderia pseudomallei.";
 RL Proc. Natl. Acad. Sci. U.S.A. 101:14240-14245(2004).
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivative license

CC -----
 CC EMBL, BX571965; CAH37269.1; -; Genomic DNA.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005279; F:amino acid-polyamine transporter activity; IEA.
 DR GO; GO:0006865; P:amino acid transport; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR002293; AA/rel_permease1.
 DR InterPro; IPR004840; AAC_permease.
 DR InterPro; IPR004841; Permease_region.
 DR PANTHER; PTHR11785; AA/rel_permease1; 1.
 DR Pfam; PF00324; AA_permease; 1.
 DR PROSITE; PS00218; AMINO_ACID_PERMEASE_1; 1.
 KM Complete proteome.
 SQ SEQUENCE 468 AA; 50875 MW; 69EB0D62E681105 CRC64;

Query Match 100.0%; Score 31; DB 2; Length 468;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRYAR 6
 |||||
 Db 78 FSRYAR 83

RESULT 20
 ID ASCB_ECOLI STANDARD; PRT; 474 AA.
 AC P24240; P78104; O59375;
 DT 01-MAR-1992, integrated into UniProtKB/Swiss-Prot.
 DT 01-NOV-1997, sequence version 2.
 DT 07-MAR-2006, entry version 51.
 DE 6-phospho-beta-glucosidase ascb (EC 3.2.1.86).
 GN Name=ascB; OrderedLocustNames=b2716;
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxId=562;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
 RC STRAIN=K12;
 RX MEDLINE=92334140; PubMed=1630307;
 RA Hall B.G., Xu L.;
 RT "Nucleotide sequence, function, activation, and evolution of the
 RT cryptic asc operon of Escherichia coli K12.";
 RL Mol. Biol. Evol. 9:688-706(1992).
 RN [2]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503; DOI=10.1126/science.277.5331.1453;
 RA Blattner F.R., Plunkett G., III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 CC -1- FUNCTION: Can hydrolyze salicin, cellobiose, and probably arbutin.
 CC -1- CATALYTIC ACTIVITY: 6-phospho-beta-D-glucosyl-(1,4)-D-glucose +
 CC H(2)O = D-glucose + D-glucose 6-phosphate.
 CC -1- SIMILARITY: Belongs to the glycosyl hydrolase 1 family.

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivative license
 CC -----
 CC EMBL; M73326; AAA16430.1; -; Unassigned DNA.
 DR EMBL; U29579; AAA69226.1; ALT INIT; Genomic DNA.
 DR EMBL; U00096; AAC75758.1; -; Genomic DNA.
 DR PIR; H65051; H65051.
 DR HSP; P11546; 1PBG.
 DR GenomeReviews; U00096_GR; b2716.
 DR Echobase; EB0083; -;
 DR Ecocyc; ECG10085; ascb.
 DR BioCyc; ECG10085-MONOMER; -;
 DR InterPro; IPR001360; Glyco_hydro_1.
 DR PANTHER; PTHR10353; Glyco_hydro_1; 1.
 DR Pfam; PF00232; Glyco_hydro_1; 1.
 DR PRINTS; PR00131; GLYHYDRLASE1.
 DR PROSITE; PS00572; GLYCOSYL_HYDROL_F1_1; 1.
 DR PROSITE; PS00653; GLYCOSYL_HYDROL_F1_2; 1.
 KM Complete proteome; Glycosidase; Hydrolase.
 FT CHAIN 1 474
 FT ACT_SITE 180 180
 FT ACT_SITE 372 372
 FT ACT_SITE 405 406
 FT CONFLICT 428 428
 FT CONFLICT 455 456
 FT CONFLICT 474 AA; 53935 MW; 02AC6BEBF211011 CRC64;
 SQ SEQUENCE 474 AA; 53935 MW; 02AC6BEBF211011 CRC64;

Query Match 100.0%; Score 31; DB 1; Length 474;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRYAR 6
 |||||
 Db 157 FSRYAR 162

```

RESULT 21
Q31X80_SHIBS PRELIMINARY; PRT; 474 AA.
ID Q31X80;
AC Q31X80;
DT 06-DEC-2005, integrated into UniProtKB/TrEMBL.
DT 06-DEC-2005, sequence version 1.
DT 07-FEB-2006, entry version 4.
DE 6-phospho-beta-glucosidase, cryptic.
GN Name=asCB; OrderedlocusNames=SB02802;
OS Shigella boydii serotype 4 (strain Sb227).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=300267;

RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP PubMed=16275786; DOI=10.1093/nar/gki954;
RX Yang F., Yang J., Zhang X., Chen L., Jiang Y., Yan Y., Tang X.,
RA Wang J., Xiong Z., Dong J., Xue Y., Zhu Y., Xu X., Sun L., Chen S.,
RA Nie H., Peng J., Xu J., Wang Y., Yuan Z., Wen Y., Yao Z., Shen Y.,
RA Qiang B., Hou Y., Yu J., Jin Q.,
RT "Genome dynamics and diversity of Shigella species, the etiologic
agents of bacillary dysentery."
RL Nucleic Acids Res. 33:6445-6458(2005).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license

DR EMBL; CP000036; AB67328.1; -; Genomic DNA.
DR GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
KM Complete proteome.
SQ SEQUENCE 474 AA; 53786 MW; 54823AE3832FF87F CRC64;

Query Match 100.0%; Score 31; DB 2; Length 474;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRYAR 6
Db 157 FSRYAR 162

RESULT 22
Q32CL4_SHIDS PRELIMINARY; PRT; 474 AA.
ID Q32CL4;
AC Q32CL4;
DT 06-DEC-2005, integrated into UniProtKB/TrEMBL.
DT 06-DEC-2005, sequence version 1.
DT 07-FEB-2006, entry version 4.
DE 6-phospho-beta-glucosidase.
GN Name=asCB; OrderedlocusNames=SDY2912;
OS Shigella dysenteriae serotype 1 (strain Sd197).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=300267;

RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP PubMed=16275786; DOI=10.1093/nar/gki954;
RX Yang F., Yang J., Zhang X., Chen L., Jiang Y., Yan Y., Tang X.,
RA Wang J., Xiong Z., Dong J., Xue Y., Zhu Y., Xu X., Sun L., Chen S.,
RA Nie H., Peng J., Xu J., Wang Y., Yuan Z., Wen Y., Yao Z., Shen Y.,
RA Qiang B., Hou Y., Yu J., Jin Q.,
RT "Genome dynamics and diversity of Shigella species, the etiologic
agents of bacillary dysentery."
RL Nucleic Acids Res. 33:6445-6458(2005).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license

DR EMBL; CP000034; AB67294.1; -; Genomic DNA.
DR GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
KM Complete proteome.

```

```

SQ SEQUENCE 474 AA; 53786 MW; A6572D302DCABCC8 CRC64;

Query Match 100.0%; Score 31; DB 2; Length 474;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRYAR 6
Db 157 FSRYAR 162

RESULT 23
Q3YER8_SHISS PRELIMINARY; PRT; 474 AA.
ID Q3YER8;
AC Q3YER8;
DT 27-SEP-2005, integrated into UniProtKB/TrEMBL.
DT 27-SEP-2005, sequence version 1.
DT 07-FEB-2006, entry version 5.
DE 6-phospho-beta-glucosidase, cryptic.
GN Name=asCB; OrderedlocusNames=SSO_2860; ORFNames=SSO_2860;
OS Shigella sonnei (strain Ss046).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=300269;

RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP PubMed=16275786; DOI=10.1093/nar/gki954;
RX Yang F., Yang J., Zhang X., Chen L., Jiang Y., Yan Y., Tang X.,
RA Wang J., Xiong Z., Dong J., Xue Y., Zhu Y., Xu X., Sun L., Chen S.,
RA Nie H., Peng J., Xu J., Wang Y., Yuan Z., Wen Y., Yao Z., Shen Y.,
RA Qiang B., Hou Y., Yu J., Jin Q.,
RT "Genome dynamics and diversity of Shigella species, the etiologic
agents of bacillary dysentery."
RL Nucleic Acids Res. 33:6445-6458(2005).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license

DR EMBL; CP000038; AA289464.1; -; Genomic DNA.
DR GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR001360; Glyco_hydro_1.
DR Pfam; PF00232; Glyco_hydro_1; 1.
DR PRINTS; PR00133; GLYHYDRLASE1.
DR PROSITE; PS00572; GLYCOSYL_HYDROL_F1_1; 1.
DR PROSITE; PS00653; GLYCOSYL_HYDROL_F1_2; 1.
KM Complete proteome.
SQ SEQUENCE 474 AA; 53830 MW; 2FC8B759849919C7 CRC64;

Query Match 100.0%; Score 31; DB 2; Length 474;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRYAR 6
Db 157 FSRYAR 162

RESULT 24
Q2MAB3_ECOLI PRELIMINARY; PRT; 474 AA.
ID Q2MAB3;
AC Q2MAB3;
DT 21-FEB-2006, integrated into UniProtKB/TrEMBL.
DT 21-FEB-2006, sequence version 1.
DT 07-MAR-2006, entry version 2.
DE Cryptic 6-phospho-beta-glucosidase.
GN Name=asCB;
OS Escherichia coli W3110.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=316407;
RN NUCLEOTIDE SEQUENCE.

```

RC STRAIN=K-12;
RX MEDLINE=81053692; PubMed=6159575;
RA Smith D.R., Calvo J.M.;
RT "Nucleotide sequence of the E coli gene coding for dihydrofolate
RT reductase.";
RL Nucleic Acids Res. 8:2255-2274(1980).
RN [12]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RA Sekiguchi T., Ortega-Cesena J., Nosch Y., Ohashi S., Tsuda K.,
RA Kanaya S.;
RT "DNA and amino-acid sequences of 3-isopropylmalate dehydrogenase of
RT Bacillus coagulans. Comparison with the enzymes of Saccharomyces
RT cerevisiae and Thermus thermophilus.";
RL Biochim. Biophys. Acta 867:36-44(1986).
RN [13]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RA Chen H., Sun Y., Stark T., Beattie W., Moses R.E.;
RT "Nucleotide sequence and deletion analysis of the polB gene of
RT Escherichia coli.";
RL DNA Cell Biol. 9:613-635(1990).
RN [14]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RA Smallshaw J.E., Kelln R.A.;
RT "Cloning, nucleotide sequence and expression of the Escherichia coli
RT K-12 pyrH gene encoding UMP kinase.";
RL Genetics (Life Sci. Adv.) 11:59-65(1992).
RN [15]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RA Hayashi K., Morooka N., Yamamoto Y., Fujita K., Isono K., Choi S.,
RA Ohtsubo E., Baba T., Wanner B.L., Mori H., Horiuchi T.;
RT "Highly accurate genome sequences of Escherichia coli K-12 strains
RT MG1655 and W3110.";
RL Mol. Syst. Biol. 0:0-0(2006).
RN [16]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RA PubMed=16397293; DOI=10.1093/nar/gk1150;
RA Riley M., Abe T., Arnaud M.B., Berlyn M.K., Blattner F.R.,
RA Chaudhuri R.R., Glaesner J.D., Horiuchi T., Keseler I.M., Koenig T.,
RA Mori H., Perna N.T., Plunkett G. III, Rudd K.E., Serres M.H.,
RA Thomas G.H., Thomson N.R., Wishart D., Wanner B.L.;
RT "Escherichia coli K-12: a cooperatively developed annotation snapshot-
RT -2005.";
RL Nucleic Acids Res. 34:1-9(2006).
RN [17]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RX MEDLINE=97251358; PubMed=9097040; DOI=10.1093/dnares/3.6.379;
RA Itoh T., Alpha H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
RA Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nishimoto H.,
RA Oshima T., Oyama S., Saito N., Sampei G., Satoh Y., Sivasundaram S.,
RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
RA Yamagata S., Horiuchi T.;
RT "Construction of a contiguous 874-kb sequence of the Escherichia coli
RT K-12 genome corresponding to 50.0-68.8 min on the linkage map and
RT analysis of its sequence features.";
RL DNA Res. 4:91-113(1997).
RN [18]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RX MEDLINE=97251358; PubMed=9097040; DOI=10.1093/dnares/3.6.379;
RA Itoh T., Alpha H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
RA Kasai H., Kimura S., Kitakawa M., Kitagawa M., Makino K., Miki T.,
RA Mizobuchi K., Mori H., Mori T., Motomura K., Nakade S., Nakamura Y.,
RA Nishimoto H., Nishio Y., Oshima T., Saito N., Sampei G., Seki Y.,
RA Sivasundaram S., Tagami H., Takeda J., Takemoto K., Wada C.,
RA Yamamoto Y., Horiuchi T.;
RT "A 460-kb DNA sequence of the Escherichia coli K-12 genome

RT corresponding to the 40.1-50.0 min region on the linkage map.";
RL DNA Res. 3:379-392(1996).
RN [19]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RX MEDLINE=97251357; PubMed=9097039; DOI=10.1093/dnares/3.6.363;
RA Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
RA Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T., Saito N.,
RA Sampei G., Seki Y., Sivasundaram S., Tagami H., Takeda J.,
RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;
RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 28.0-40.1 min region on the linkage map.";
RL DNA Res. 3:363-377(1996).
RN [10]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RX MEDLINE=97034878; PubMed=8940112; DOI=10.1074/jbc.271.49.31145;
RA Arn E.A., Abelson J.N.;
RT "The 2'-5' RNA ligase of Escherichia coli. Purification, cloning, and
RT genomic disruption.";
RL J. Biol. Chem. 271:31145-31153(1996).
RN [11]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RX MEDLINE=97061202; PubMed=8905232; DOI=10.1093/dnares/3.3.137;
RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Horjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 12.7-28.0 min region on the linkage map.";
RL DNA Res. 3:137-155(1996).
RN [12]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RX MEDLINE=94261430; PubMed=8202364;
RA Fujita N., Mori H., Yura T., Ishihama A.;
RT "Systematic sequencing of the Escherichia coli genome: analysis of the
RT 2.4-4.1 min (110,917-193,643 bp) region.";
RL Nucleic Acids Res. 22:1637-1639(1994).
RN [13]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RX MEDLINE=94240115; PubMed=8189897;
RA Janosi L., Shimizu I., Kaji A.;
RT "Ribosome recycling factor (ribosome releasing factor) is essential
RT for bacterial growth.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:4249-4253(1994).
RN [14]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RX MEDLINE=94124004; PubMed=7904973; DOI=10.1016/0378-1119(93)90470-N;
RA Allikmets R., Gerrard B.C., Court D., Dean M.C.;
RT "Cloning and organization of the abc and mdl genes of Escherichia
RT coli: relationship to eukaryotic multidrug resistance.";
RL Gene 136:231-236(1993).
RN [15]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RX MEDLINE=94018640; PubMed=8412694;
RA van Heeswijk W.C., Rabenberg M., Westerhoff H.V., Kahn D.D.;
RT "The genes of the gluconate synthetase adenyllylation cascade are not
RT regulated by nitrogen in Escherichia coli.";
RL Mol. Microbiol. 9:443-458(1993).
RN [16]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RX MEDLINE=93259920; PubMed=8387990;
RA Zhao S., Sandt C.H., Feulner G., Vlazny D.A., Gray J.A., Hill C.W.;


```

RT "Rhs elements of Escherichia coli K-12: complex composites of shared
RT and unique components that have different evolutionary histories.";
RL J. Bacteriol. 175:2799-2808(1993).
RN (17)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RX MEDLINE=93123180; PubMed=8419307;
RA Yamada M., Asaoka S., Saiter M.H. Jr., Yamada Y.;
RT "Characterization of the gsd gene from Escherichia coli K-12 W3110 and
RT regulation of its expression.";
RL J. Bacteriol. 175:568-571(1993).
RN (18)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RX MEDLINE=93116053; PubMed=1474579;
RA Cornack R.S., Mackie G.A.;
RT "Structural requirements for the processing of Escherichia coli 5 S
RT ribosomal RNA by RNase E in vitro.";
RL J. Mol. Biol. 228:1078-1090(1992).
RN (19)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RX MEDLINE=93094132; PubMed=1459951;
RA Gervais F.G., Drapeau G.R.;
RT "Identification, cloning, and characterization of rcsF, a new
RT regulator gene for exopolysaccharide synthesis that suppresses the
RT division mutation ftsZ84 in Escherichia coli K-12.";
RL J. Bacteriol. 174:8015-8022(1992).
RN (20)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RX MEDLINE=93077430; PubMed=1447125;
RA Yamataka K., Ogura T., Niki H., Hiraga S.;
RT "Identification and characterization of the smbA gene, a suppressor of
RT the mukB null mutant of Escherichia coli.";
RL J. Bacteriol. 174:7517-7526(1992).
RN (21)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RX MEDLINE=93011013; PubMed=1396599;
RA Condon C., Phillips J., Fu Z.Y., Squires C., Squires C.L.;

```

```

Query Match 100.0%; Score 31; DB 2; Length 474;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 FSRYAR 6
DB 157 FSRYAR 162

```

```

RESULT 25
O8X841 EC057
AC O8X841 EC057 PRELIMINARY; PRT; 474 AA.
DT 01-MAR-2002. Integrated into UniProtKB/TrEMBL.
DT 01-MAR-2002. Sequence version 1.
DE 6-phospho-beta-glucosidase; entry version 26.
DE 6-phospho-beta-glucosidase; cryptic (6-phospho-beta-glucosidase).
GN NameasGB; OrderedlocusNames=EC63572, z4024;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN (1)
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
RA Berna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Posei D.V., Wayne G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Rosati G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimailanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,

```

```

RA Welch R.A., Blatner F.R.;
RT "Genome sequence of enterohemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RN (2)
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O157:H7 / Sakai / RIMD 0509952 / EHEC;
RX MEDLINE=21156231; PubMed=11258796; DOI=10.1093/nar/29/8.1.11;
RA Hayashi T., Makino K., Onishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hatton M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).

```

```

CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
EMBL; AB005174; AAG57823.1; -; Genomic DNA.
DR EMBL; BA000007; BAB36995.1; -; Genomic DNA.
DR PIR; C85920; C85920.
DR PIR; D91075; D91075.
DR HSSP; P11546; 1PBG.
DR BIOCyc; EC048334-1:EC53572-MONOMER; -.
DR GO; GO:0004553; F.hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.
DR GO; GO:0005975; P.carbohydrate metabolism, IEA.
DR INTERPro; IPR001360; Glyco_hydro_1.
DR PANTHER; PTHR10353; Glyco_hydro_1; 1.
DR Pfam; PF00232; Glyco_hydro_1; 1.
DR PRINTS; PR00131; GLYDRLASE1.
DR PROSITE; PS00572; GLYCOSYL_HYDROL_F1_1; 1.
DR PROSITE; PS00653; GLYCOSYL_HYDROL_F1_2; 1.
KW Complete proteome.
SQ
SEQUENCE 474 AA; 53731 MW; 7B31E492C9ABD14 CRC64;

```

```

Query Match 100.0%; Score 31; DB 2; Length 474;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 FSRYAR 6
DB 157 FSRYAR 162

```

```

RESULT 26
O3XAM6 PSEPP PRELIMINARY; PRT; 494 AA.
AC O3XAM6;
DT 08-NOV-2005. Integrated into UniProtKB/TrEMBL.
DT 21-FEB-2006. Sequence version 1.
DE Permease for cytosine/purines, uracil, thiamine, allantoin.
GN OrderedlocusNames=Pfl_3440;
OS Pseudomonas fluorescens (strain Pfl-1).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=205922;
RN (1)
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RG US DOE Joint Genome Institute;
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler J.C., Glavina T.,
RA Hammon N., Israni S., Plunkett S., Saunders E.H., Schmutz J.,
RA Larrimer F., Land M., Kyriakides N., Anderson I., Richardson P.;
RT "Complete sequence of Pseudomonas fluorescens Pfl-1.";
RL Submitted (AUG-2005) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
EMBL; CP000094; ABA75178.1; -; Genomic DNA.
DR GO; GO:0016020; C.membrane; IEA.
DR GO; GO:0015205; F.nucleobase transporter activity; IEA.
DR GO; GO:0015931; P.nucleobase, nucleoside, nucleotide and nucl. . .; IEA.

```

```

DR InterPro: IPR001248; Cyt_pur_permease.
DR Pfam: PF02133; Transp_cyt_pur; 1.
DR Complete proteome.
SQ SEQUENCE 494 AA; 53346 MW; 593C147D6A2D89FD CRC64;

Query Match
Best Local Similarity 100.0%; Score 31; DB 2; Length 494;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSRYAR 6
Db 257 FSRYAR 262

RESULT 27
Q2X8L3_PSEPU PRELIMINARY; PRT; 496 AA.
AC Q2X8L3_
DT 10-JAN-2006, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, entry version 1.
DE Transporter, NCS1 nucleoside transporter family.
GN ORFNames=PputDRAFT_0203;
OS Pseudomonas putida_F1
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=351746;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=F1;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Pseudomonas putida F1."
RT "Sequencing of the draft genome and assembly of Pseudomonas putida F1."
RT Submitted (OCT-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=F1;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Pseudomonas putida F1."
RT Submitted (NOV-2005) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is preliminary data.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
DR EMBL; AALM01000087; EAP48124.1; -; Genomic_DNA.
DR SEQUENCE 496 AA; 53548 MW; 4681791B53DFB357 CRC64;

Query Match
Best Local Similarity 100.0%; Score 31; DB 2; Length 496;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSRYAR 6
Db 257 FSRYAR 262

RESULT 28
Q88FQ2_PSEPK PRELIMINARY; PRT; 496 AA.
AC Q88FQ2_
DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2003, sequence version 1.
DE Transporter, NCS1 nucleoside transporter family.
GN OrderedLocusNames=PP4035; ORFNames=PP_4035;
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;

```

```

OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=22423060; PubMed=12534463;
RX DOI=10.1046/j.1462-2920.2002.00366.x;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L.M., Beaman M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,
RA Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,
RA Hance I., Chris Lee P., Holtzaple E.K., Scanlan D., Tran K.,
RA Moazzes A., Uterback T.R., Rizzo M., Lee K., Kosack D., Moestl D.,
RA Wedler H., Lauber J., Stegandic D., Hohnselt J., Straczek M., Helm S.,
RA Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoelt A., Tuenmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the RT metabolically versatile Pseudomonas putida KT2440."
RL Environ. Microbiol. 4:799-808(2002).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
DR EMBL; AE015451; AAN69627.1; -; Genomic_DNA.
DR TIGR; PP4035;
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0015205; F:nucleoside transporter activity; IEA.
DR GO; GO:0015931; P:nucleoside, nucleoside, nucleotide and nucl. .; IEA.
DR InterPro: IPR001248; Cyt_pur_permease.
DR Pfam: PF02133; Transp_cyt_pur; 1.
KW Complete proteome.
SQ SEQUENCE 496 AA; 53520 MW; 4681791B4CCFB357 CRC64;

Query Match
Best Local Similarity 100.0%; Score 31; DB 2; Length 496;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSRYAR 6
Db 257 FSRYAR 262

RESULT 29
Q91674_PSEAE PRELIMINARY; PRT; 496 AA.
AC Q91674_
DT 01-MAR-2001, integrated into UniProtKB/TrEMBL.
DT 01-MAR-2001, sequence version 1.
DE Transporter, NCS1 nucleoside transporter family.
DE Probable transporter.
GN OrderedLocusNames=PA0443;
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=ATCC 15692 / PAO1;
RX MEDLINE=20437337; PubMed=10984043; DOI=10.1038/35023079;
RA Stover C.K., Pham X.-O.T., Erwin A.L., Mitsuuchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.U., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H. Jr., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathogen."
RT Nature 406:959-964(2000).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
DR EMBL; AE004481; AAG03832.1; -; Genomic_DNA.
DR FTR; B83591; B83591.

```

```

DR BiOCyc: PA0443-MONOMER; -.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0015205; F:nucleobase transporter activity; IEA.
DR GO: GO:0015931; F:nucleobase, nucleoside, nucleotide and nucl. .; IEA.
DR InterPro: IPR001248; Cyt_pur_permease.
DR Pfam: PF02133; Transp_cyt_pur; 1.
KM Complete proteome.
SQ SEQUENCE 496 AA; 53681 MW; C9DAC04028E2D4C7 CRC64;

Query Match 100.0%; Score 31; DB 2; Length 496;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRYPAR 6
DB 257 FSRYPAR 262

RESULT 30
Q3BYW8_XANCS PRELIMINARY; PRT; 526 AA.
ID Q3BYW8_XANCS
AC Q3BYW8;
DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.
DT 22-NOV-2005, sequence version 1.
DT 07-FEB-2006, entry version 5.
DE Gamma-glutamyltranspeptidase (EC 2.3.2.2).
GN Name=ggf1; OrderedLocNames=XCV0314;
OS Xanthomonas campestris pv. vesicatoria (strain 85-10).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=316273;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=16237009; DOI=10.1126/JB.187.21.7254-7266.2005;
RA Thieme F., Koehnig R., Bekel T., Berger C., Boch J., Buettner D.,
RA Caldana C., Gaigalat U., Goemann A., Kay S., Kitchner O., Lanz C.,
RA Linke B., McHardy A.C., Meyer F., Mittenhuber G., Nies D.H.,
RA Niesbach-Kloeppen U., Patschkowski T., Rueckert C., Rupp O.,
RA Schneider S., Schuster S.C., Vorhoefer F.J., Weber E., Puhler A.,
RA Bonas U., Bartels D., Kaiser O.;
RT "Insights into genome plasticity and pathogenicity of the plant
RT pathogenic bacterium Xanthomonas campestris pv. vesicatoria revealed
RT by the complete genome sequence.";
RL J. Bacteriol. 187:7254-7266(2005).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
DR EMBL: AM039952; CAJ21945.1; -; Genomic DNA.
DR GO: GO:0008415; F:acyltransferase activity; IEA.
DR GO: GO:0003840; F:gamma-glutamyltransferase activity; IEA.
DR GO: GO:0016740; F:transferase activity; IEA.
KM Acyltransferase; Complete proteome; Transferase.
SQ SEQUENCE 526 AA; 55499 MW; B40B3B743818AA43 CRC64;

Query Match 100.0%; Score 31; DB 2; Length 526;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRYPAR 6
DB 434 FSRYPAR 439

RESULT 31
Q8POL8_XANAC PRELIMINARY; PRT; 526 AA.
ID Q8POL8_XANAC
AC Q8POL8;
DT 01-OCT-2002, integrated into UniProtKB/TrEMBL.
DT 01-OCT-2002, sequence version 1.
DT 07-FEB-2006, entry version 15.
DE Gamma-glutamyltranspeptidase.
GN Name=ggf;

```

```

OS Xanthomonas axonopodis pv. citri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=22022145; PubMed=12024217; DOI=10.1038/417459a;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A.,
RA Almeida N.F., Jr., Alves L.M.C., do Amaral A.M., Bertolini M.C.,
RA Camargo L.E.A., Camarotte G., Cannavan F., Cardozo J., Chambergro F.,
RA Ciapina L.P., Ciccarelli R.M.B., Coutinho L.L., Cursino-Santos J.R.,
RA El-Dorri H., Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Gruber A.,
RA Ferro M.I.T., Fornighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spindola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitzajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities.";
RL Nature 417:459-463(2002).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
DR EMBL: AE011655; AAM35197.1; -; Genomic DNA.
DR MEROPS; T03.014; -.
DR LinkHub; Q8POL8; -.
DR GO: GO:0003840; F:gamma-glutamyltransferase activity; IEA.
DR InterPro: IPR001011; GGT_peptidase.
DR PANTHER: PTHR11686; GGT_peptidase; 1.
DR Pfam: PF01019; G_glu_transpept; 1.
DR PRINTS: PR01210; GGTTRANSPTASE.
DR TIGRFAMs: TIGR00066; g_glu_trans; 1.
KM Complete proteome.
SQ SEQUENCE 526 AA; 55544 MW; C841558896A9B64 CRC64;

Query Match 100.0%; Score 31; DB 2; Length 526;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRYPAR 6
DB 434 FSRYPAR 439

RESULT 32
Q7SA29_NEUCR PRELIMINARY; PRT; 560 AA.
ID Q7SA29_NEUCR
AC Q7SA29;
DT 15-DEC-2003, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, entry version 11.
DE Hypothetical protein.
GN ORFNames=NCU07334.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=74-OR23-1A / FGSC 987;
RX MEDLINE=22598136; PubMed=12712197; DOI=10.1038/nature01554;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., Fitzhugh W., Ma L.-J., Smirnov S., Futrell S., Rehman B.,
RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Iankiev P., Bell-Pedersen D., Nelson M.A.,
RA Werner-Washburne M., Selitrennikoff C.P., Kinsey J.A., Braun E.L.,

```

RA Zelter A., Schulte U., Kothe G.O., Jedd G., Nemes H.-W., Staben C.,
 RA Marcotte E., Greenberg D., Roy A., Foley K., Naylor J.,
 RA Stange-Thomann N., Barrett R., Gierre S., Kamal M., Kamysseilis M.,
 RA Muscili E., Bielke C., Rudd S., Frishman D., Krystofova S.,
 RA Rasmussen C., Metznerberg R.L., Perkins D.D., Kroken S., Cogoni C.,
 RA Macino G., Catcheside D.E.A., Li W., Pratt R.J., Osman S.A.,
 RA Desouza C.P.C., Glass N.L., Orbach M.J., Berglund J.A., Voelker R.,
 RA Yarden O., Plamann M., Sella S., Dunlap J.C., Radford A., Aramayo R.,
 RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,
 RA Paulsen I., Sachs M.S., Lander E.S., Nisbaum C., Birren B.W.,
 RT "The genome sequence of the filamentous fungus *Neurospora crassa*,"
 RT Nature 422:859-868(2003).
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs License
 CC -----
 DR EMBL: AABX01000207; EAA33256.1; -; Genomic_DNA.
 DR GO: GO:0016020; C:membrane; IEA.
 DR GO: GO:0015205; F:nucleobase transporter activity; IEA.
 DR GO: GO:0015931; P:nucleobase, nucleoside, nucleotide and nucl. . .; IEA.
 DR InterPro: IPR001248; Cyt_pur_permease.
 DR InterPro: IPR012681; NCS1.
 DR Pfam: PF02133; Transp_cyt_pur; 1.
 DR TIGRFAMs: TIGR00800; ncs1; 1.
 KM Hypothetical protein.
 SQ SEQUENCE 560 AA; 62325 MW; FAC13B1A5F26565 CRC64;

Query Match 100.0%; Score 31; DB 2; Length 560;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSRYAR 6
 Db 269 FSRYAR 274

RESULT 33
 Q21282_RHOA PRELIMINARY; PRT; 635 AA.
 AC Q21282;
 DT 07-MAR-2006, integrated into UniProtKB/TrEMBL.
 DT 07-MAR-2006, sequence version 1.
 DE Extracellular solute-binding protein, family 5 precursor.
 GN ORFNames=RPB_1770;
 OS Rhodospseudomonas palustris Haaz2.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Bradyrhizobiaceae; Rhodospseudomonas.
 OX NCBI_TaxID=316058;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=Haaz2;
 RA US DOE Joint Genome Institute;
 RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler J.C., Glavina T.,
 RA Hammon N., Israni S., Pliukus S., Chain P., Malfacti S., Shin M.,
 RA Vergez L., Schmutz J., Larimer F., Land M., Hauser L., Pelletier D.A.,
 RA Kyriakides N., Anderson I., Oda Y., Harwood C.S., Richardson P.,
 RT "Complete sequence of *Rhodospseudomonas palustris* Haaz2,"
 RT Submitted (JAN-2006) to the EMBL/GenBank/DBJ databases.
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs License
 CC -----
 DR EMBL: CP000250; ABD06478.1; -; Genomic_DNA.
 KM Signal.
 FT SIGNAL
 SQ SEQUENCE 635 AA; 71469 MW; 08D975473193E710 CRC64;
 Query Match 100.0%; Score 31; DB 2; Length 635;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 FSRYAR 6
 Db 595 FSRYAR 600

RESULT 34
 Q374J7_RHOA PRELIMINARY; PRT; 663 AA.
 ID Q374J7_RHOA
 AC Q374J7;
 DT 06-DEC-2005, integrated into UniProtKB/TrEMBL.
 DT 06-DEC-2005, sequence version 1.
 DE Twin-arginine translocation pathway signal precursor.
 GN ORFNames=RPEDRAFT_2330;
 OS Rhodospseudomonas palustris BisA53.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Bradyrhizobiaceae; Rhodospseudomonas.
 OX NCBI_TaxID=316055;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=BisA53;
 RA US DOE Joint Genome Institute (JGI-ORNL);
 RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler J.C., Glavina T.,
 RA Hammon N., Israni S., Pliukus S., Richardson P.,
 RT "Sequencing of the draft genome and assembly of *Rhodospseudomonas*
 RT palustris BisA53,"
 RT Submitted (OCT-2005) to the EMBL/GenBank/DBJ databases.
 RN [2]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=BisA53;
 RA US DOE Joint Genome Institute (JGI-ORNL);
 RA Larimer F., Land M.,
 RT "Annotation of the draft genome of *Rhodospseudomonas palustris*
 RT BisA53," (OCT-2005) to the EMBL/GenBank/DBJ databases.
 RL Submitted (OCT-2005) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs License
 CC -----
 DR EMBL: AALA01000008; EAO90388.1; -; Genomic_DNA.
 DR GO: GO:0005215; F:transporter activity; IEA.
 DR GO: GO:0006810; P:transport; IEA.
 KM Signal.
 FT SIGNAL
 SQ SEQUENCE 663 AA; 74262 MW; 32F26A217EC705E4 CRC64;

Query Match 100.0%; Score 31; DB 2; Length 663;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSRYAR 6
 Db 622 FSRYAR 627

RESULT 35
 Q4SCJ9_TETNG PRELIMINARY; PRT; 809 AA.
 ID Q4SCJ9_TETNG
 AC Q4SCJ9;
 DT 19-JUL-2005, integrated into UniProtKB/TrEMBL.
 DT 19-JUL-2005, sequence version 1.
 DE Chromosome undetermined SCAP14653, whole genome shotgun sequence.
 DE (Fragment).
 GN ORFNames=GSTEING0020461001;
 OS Tetradon nigroviridis (Green puffer).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 OC Tetraodontidae; Tetraodontidae; Tetraodon.
 RN NCBI_TaxID=99883;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX PubMed=15496914; DOI=10.1038/nature03025;
 RA Jallion O., Aury J.-M., Brunet P., Petit J.-L., Stange-Thomann N.,
 Mauceli E., Bonneau L., Fischer C., Ouzif-Costaz C., Bernot A.,
 Nicard S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
 Dastiva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
 Anthouard V., Jubin C., Castell V., Katinka M., Vacherie B.,
 Biemont C., Skali Z., Catolico L., Poulain J., De Berardinis V.,
 Cruaud C., Duprat S., Brottier P., Coutanceau J.-P., Gouzy J.,
 Parra G., Lardier G., Chapple C., McKernan K.J., McMan P., Bosak S.,
 Kellis M., Wolff J.-N., Guigo R., Zody M.C., Mesirov J.,
 Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
 Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
 Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.,
 RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
 the early vertebrate proto-karyotype".
 RL Nature 431:946-957(2004).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RG Genoscope; Whitehead Institute Centre for Genome Research;
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- CAUTION: The sequence shown here is derived from an
 EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 preliminary data.
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC distributed under the Creative Commons Attribution-NonDerivative license
 CC -----
 DR EMBL: CAAB01014653; CAG01633.1; -, Genomic DNA.
 DR GO: GO:0003676; F:nucleic acid binding; IEA.
 DR GO: GO:0006869; P:lipid transport; IEA.
 DR GO: GO:0008202; P:steroid metabolism; IEA.
 DR InterPro: IPR000008; C2.
 DR InterPro: IPR000648; Oxyesterol_bd.
 DR InterPro: IPR001849; PH.
 DR PANTHER: PTHR10972; Oxyesterol_BP; 1.
 DR Pfam: PF01237; Oxyesterol_BP; 1.
 DR PRINTS: PR00360; C2DOMAIN.
 DR PROSITE: PS01013; OSBP; 1.
 DR PROSITE: PS50003; PH_DOMAIN; 1.
 KW lipid transport; Transport.
 FT NON TER 809 809
 SQ SEQUENCE 809 AA; 91216 MW; 3199FOA349871EB4 CRC64;
 Query Match 100.0%; Score 31; DB 2; Length 809;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FSRYAR 6
 DB 742 FSRYAR 747

RA Birren B.W., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
 RA Alt-Zahra M., Allen R., Allen T., An P., Anderson M., Anderson S.,
 RA Arachchi H.M., Armbruster J., Bachantang P., Baldwin J., Barry A.,
 RA Bayut T., Blichstein B., Bloom T., Biye J., Boguslavsky L.,
 RA Borowsky M., Boukhalil B., Brunache A., Butler J., Calixte N.,
 RA Calvo S.E., Camarata J., Campo K., Chang K., Cheshatsang Y.,
 RA Citeron M., Collamore A., Considine T., Cook A., Cooke P., Corum B.,
 RA Cuomo C., David R., Dawoe T., Degray S., Dodge S., Doolley K.,
 RA Dorje P., Dorje K., Doris L., Dufey N., Dupes A., Elkins T.,
 RA Dorje P., Erickson J., Farina A., Fato S., Ferreira P., Fischer H.,
 RA Fitzgerald M., Foley A., Gage D., Galagan J.B., Geatin G., Gierre S.,
 RA Glatke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
 RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
 RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husbey E., Iliev I.,
 RA Jaffe D., Jones C., Kamal M., Kamat A., Kanyseels M., Karlsson E.,
 RA Kells C., Kieu A., Kiser P., Kodira C., Kulbokas E., Labutti K.,
 RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
 RA Lindblad-Toh K., Liu X., Lokysang T., Lokysang Y., Lucien O.,
 RA Lui A., Ma L.-d., Mabbitt R., MacDonald J., Maclean C., Major J.,
 RA Manning J., Maraballa R., Maru K., Matthews C., Mauceli E.,
 RA McCarthy M., McDonough S., McInnes T., Meldrum J., Menes L.,
 RA Mesirov J., Minalev A., Minova T., Mikkelsen T., Menga V., Moru K.,
 RA Mozes J., Mulrain L., Munson G., Naylor J., News C., Nguyen C.,
 RA Nguyen N., Nguyen T., Nicol R., Nielsen C.B., Nizzari M., Norbu C.,
 RA Norbu N., O'Donnell P., Okoawo O., O'Leary S., Omotosho B.,
 RA O'Neill K., Osman S., Parker S., Perrin D., Phunkhang P., Pigani B.,
 RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond B.,
 RA Rella R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
 RA Rutman M., Schupbach R., Seaman C., Settipalli S., Sharpe T.,
 RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,
 RA Spencer B., Stalker J., Stange-Thomann N., Stavropoulos S.,
 RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuina P.,
 RA Tenzing P., Tesfaye S., Theodore J., Thoultsang Y., Topham K.,
 RA Towey S., Tsamila T., Tsomo N., Vallee D., Vassiliou H.,
 RA Venkataswamy V.S., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
 RA Wangdi T., Whitaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
 RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembek L.,
 RA Zimmer A., Zody M., Lander E.S.,
 RT "The genome sequence of *Ustilago maydis*".
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- CAUTION: The sequence shown here is derived from an
 EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 preliminary data.
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC distributed under the Creative Commons Attribution-NonDerivative license
 CC -----
 DR EMBL: AACP01000149; EAK85206.1; -, Genomic DNA.
 DR GO: GO:0004197; F:cysteine-type endopeptidase activity; IEA.
 DR GO: GO:0004221; F:ubiquitin thiolesterase activity; IEA.
 DR GO: GO:0006512; P:ubiquitin cycle; IEA.
 DR GO: GO:0006511; P:ubiquitin-dependent protein catabolism; IEA.
 DR InterPro: IPR006035; Arg_agm_form.
 DR InterPro: IPR006615; Pept_C19_N1.
 DR InterPro: IPR001394; Peptidase_C19.
 DR Pfam: PF00443; UCH; 1.
 DR SMART: SM00695; DUSP; 1.
 DR PROSITE: PS00147; ARGINASE_1; UNKNOWN_1.
 DR PROSITE: PS00972; UCH_2_1; 1.
 DR PROSITE: PS00973; UCH_2_2; 1.
 DR PROSITE: PS50235; UCH_2_3; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 1427 AA; 156945 MW; BBB060C165811664 CRC64;
 Query Match 100.0%; Score 31; DB 2; Length 1427;
 Best Local Similarity 100.0%; Pred. No. 4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FSRYAR 6
 DB 896 FSRYAR 901

```

RESULT 37
Q57UM9_9TRYRYP PRELIMINARY; PRT; 1649 AA.
ID 057UM9_9TRYRYP PRELIMINARY; PRT; 1649 AA.
AC 057UM9_9TRYRYP PRELIMINARY; PRT; 1649 AA.
DT 10-MAY-2005, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, entry version 1.
DE Mitochondrial DNA polymerase I protein C.
GN ORFNames=TB927.7.3990;
OS Trypanosoma brucei.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxId=5691;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GUTa10.1;
RA Ghedin E., Blandin G., Bartholomeu D., Caler E., Haas B., Hannick L.,
  Johnson J., Jones K., Koo H.L., Larkin C., Pal G., Peterson J.,
  Ralston J., Salzman S., Simpson A.J., Tallon L., Van Aken S.,
  Wanless D., White O., Wortman J., Fraser C.M., El-Sayed N.M.A.;
  Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
EMBL: AC159450; AAX70690.1; -; Genomic_DNA.
DR GO: GO:0003677; F:DNA binding; IEA.
DR GO: GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
DR GO: GO:0006260; P:DNA replication; IEA.
DR InterPro: IPR001098; DNA_pol.
DR InterPro: IPR002298; DNA_pol.
DR Pfam: PF00476; DNA_pol_A; 1.
DR PRINTS: PR00868; DNAPOLI.
DR PROSITE: PS00447; DNA_POLYMERASE_A; UNKNOWN 1.
SQ SEQUENCE 1649 AA; 182029 MW; 2FC69472CC2A1A1E CRC64;

Query Match
Best local Similarity 100.0%; Score 31; DB 2; Length 1649;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRYP 6
DB 5 FSRYP 10

RESULT 38
Q8MWB3_9TRYRYP PRELIMINARY; PRT; 1649 AA.
ID 08MWB3_9TRYRYP PRELIMINARY; PRT; 1649 AA.
AC 08MWB3_9TRYRYP PRELIMINARY; PRT; 1649 AA.
DT 01-OCT-2002, integrated into UniProtKB/TrEMBL.
DT 01-OCT-2002, sequence version 1.
DT 07-FEB-2006, entry version 15.
DE DNA polymerase I-like protein C.
GN Name=POLIC;
OS Trypanosoma brucei.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxId=5691;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=TREU927;
RA MEDLINE=22146561; PubMed=12150917; DOI=10.1016/S1097-2765(02)00571-3;
RT Klingbeil M.M., Motyka S.A., Englund P.T.;
  "Multiple mitochondrial DNA polymerases in Trypanosoma brucei.";
  Mol. Cell 10:1175-186(2002).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
EMBL: AF445378; AAM81964.1; -; Genomic_DNA.
DR GO: GO:0003677; F:DNA binding; IEA.
DR GO: GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
DR GO: GO:0006260; P:DNA replication; IEA.
DR InterPro: IPR001098; DNA_pol.

```

```

DR InterPro: IPR002298; DNA_pol.
DR Pfam: PF00476; DNA_pol_A; 1.
DR PRINTS: PR00868; DNAPOLI.
DR PROSITE: PS00447; DNA_POLYMERASE_A; UNKNOWN 1.
SQ SEQUENCE 1649 AA; 182043 MW; 2AABC43C737A3BB0 CRC64;

Query Match
Best local Similarity 100.0%; Score 31; DB 2; Length 1649;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRYP 6
DB 5 FSRYP 10

RESULT 39
Q49100_MYCCA PRELIMINARY; PRT; 37 AA.
ID Q49100_MYCCA PRELIMINARY; PRT; 37 AA.
AC Q49100_MYCCA PRELIMINARY; PRT; 37 AA.
DT 01-NOV-1996, integrated into UniProtKB/TrEMBL.
DT 01-NOV-1996, sequence version 1.
DT 07-FEB-2006, entry version 23.
DE DNA gyrase subunit A (Fragment).
GN Name=gyrA;
OS Mycoplasma capricolum.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxId=2095;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC27343;
RX MEDLINE=95129856; PubMed=7828871; DOI=10.1016/0378-1119(94)90653-X;
RA Sano K., Miyata M.;
RT "The gyrB gene lies opposite from the replication origin on the
  RT circular chromosome of Mycoplasma capricolum.";
  Gene 151:181-183(1994).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC27343;
RA Sano K., Miyata M.;
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC27343;
RA Sano K.-I.;
RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
EMBL: D26016; BAA05032.1; -; Genomic_DNA.
DR NON_TER 37
FT SEQUENCE 37 AA; 4428 MW; 0C83911EFCE84697 CRC64;

Query Match
Best local Similarity 90.3%; Score 28; DB 2; Length 37;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRYP 6
DB 20 FSRYP 25

RESULT 40
Q56EC7_9CAUD PRELIMINARY; PRT; 102 AA.
ID Q56EC7_9CAUD PRELIMINARY; PRT; 102 AA.
AC Q56EC7_9CAUD PRELIMINARY; PRT; 102 AA.
DT 10-MAY-2005, integrated into UniProtKB/TrEMBL.
DT 10-MAY-2005, sequence version 1.
DT 07-FEB-2006, entry version 2.
DE Hypothetical protein PHG310R237C.
GN Name=PHG310R237C; ORFNames=PHG31p234;
OS Aeromonas phage 31.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.

```

OX NCBI_TaxID=321023;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Nolan J.M., Petrov V., Bertrand C., Kirsch H.M., Karam J.D.;
 RT "Comparative analysis of the *Aeromonas bacteriophage* 31 genome."
 RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs License
 CC -----
 CC EMBL: AY962392; AAX63723.1; -; Genomic_DNA.
 DR Hypoetical protein.
 KM SEQUENCE 102 AA; 11960 MW; B591B2BF81A9A58E CRC64;
 SQ
 Query Match 90.3%; Score 28; DB 2; Length 102;
 Best Local Similarity 83.3%; Pred. No. 1.3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FSRYAR 6
 Db 68 FNRYYAR 73
 RESULT 41
 OSU3N8_BRARE PRELIMINARY; PRT; 110 AA.
 ID OSU3N8;
 AC OSU3N8;
 DT 07-DEC-2004, integrated into UniProtKB/TrEMBL.
 DT 07-DEC-2004, sequence version 1.
 DT 07-FEB-2006, entry version 10.
 DE Hypoetical protein.
 GN Name=wrp;
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 NC NCBI_TaxID=7955;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Larvae;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
 RA Diatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loughran N.A., Peters G.J., Abramson R.D., Mullen S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Wooley K.C., Hale S., Garcia A.M., Gay L.J., Hultk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzyzanski M.I., Skalska U., Smalins D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Larvae;
 RA Director MGC Project;
 RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs License
 CC -----
 CC EMBL: BC085456; AAH85456.1; -; mRNA.
 DR Ensembl: ENSDARG0000024641; Danio rerio.
 ZFIN: ZDB-GENE-030131-7696; wrp.

DR InterPro: IPR007514; CHD5.
 DR Pfam: PF04420; CHD5. 1.
 KM Hypoetical protein.
 SQ SEQUENCE 110 AA; 12330 MW; 2496B215D029C793 CRC64;
 QY 1 FSRYAR 6
 Db 5 FARYAR 10
 RESULT 42
 OSZJM2_CHICK PRELIMINARY; PRT; 126 AA.
 ID OSZJM2;
 AC OSZJM2;
 DT 23-NOV-2004, integrated into UniProtKB/TrEMBL.
 DT 23-NOV-2004, sequence version 1.
 DT 07-FEB-2006, entry version 6.
 DE Hypoetical protein.
 GN ORFNames=RCJMB04.17c17;
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 NC NCBI_TaxID=9031;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=CB; TISSUE=Bursa;
 RA Caldwell R.B., Kierzek A.M., Arakawa H., Bezzubov Y., Zaim J.,
 RA Friedler P., Kuter S., Biagodatki A., Kostovska D., Koter M.,
 RA Plachy J., Carninci P., Hayashizaki Y., Buerstedde J.M.;
 RT "Full-length cDNAs from chicken bursa lymphocytes to facilitate
 RT gene function analysis";
 RL Genome Biol. 6:R6-R6(2005).
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs License
 CC -----
 CC EMBL: AJ720412; CAG32071.1; -; mRNA.
 DR InterPro: IPR007514; CHD5.
 DR Pfam: PF04420; CHD5. 1.
 KM Hypoetical protein.
 SQ SEQUENCE 126 AA; 14426 MW; 30A1186080FEB32B CRC64;
 QY 1 FSRYAR 6
 Db 21 FARYAR 26
 RESULT 43
 OSJUB8_BURP1 PRELIMINARY; PRT; 133 AA.
 ID OSJUB8_BURP1
 AC OSJUB8;
 DT 08-NOV-2005, integrated into UniProtKB/TrEMBL.
 DT 08-NOV-2005, sequence version 1.
 DT 21-FEB-2006, entry version 4.
 DE Hypoetical protein.
 GN OrderedLocNames=BURPS1710b_2850;
 OS Burkholderia pseudomallei (strain 1710b).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Burkholderiaceae; Burkholderia; pseudomallei group.
 NC NCBI_TaxID=320372;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RA Woods D.E., Nieman W.C.;
 RL Submitted (SEP-2005) to the EMBL/GenBank/DBJ databases.

```
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
DR EMBL: CP000124; AB48582.1; -; Genomic_DNA.
DR TIGR: B0RPS1710D_2850; -.
DR InterPro: IPR011633; DUF1602.
DR Pfam: PF07673; DUF1602; 1.
KM Complete proteome; Hypothetical protein.
SQ SEQUENCE 133 AA; 14144 MW; C069DFC776D45162 CRC64;

Query Match          90.3%; Score 28; DB 2; Length 133;
Best Local Similarity 83.3%; Pred. No. 1.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSRYAR 6
Db 80 FSRYSR 85

RESULT 44
ID Q52KS6_XENLA PRELIMINARY; PRT; 170 AA.
AC Q52KS6;
DT 24-MAY-2005, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, entry version 1.
DE Hypothetical protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodidae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Eye;
RX MEDLINE=22341133; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strusberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT Initiative";
RL Dev. Dyn. 225:384-391(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Eye;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haileh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Vallatou D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buttefield Y.S.N., Krzywinski M.I., Skalska U., Smallue D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Eye;
RA Klein S., Gerhard D.S.;
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
```

```
CC -----
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
DR EMBL: BC094207; AAH94207.1; -; mRNA.
DR Hypothetical protein.
KM Hypothetical protein.
SQ SEQUENCE 170 AA; 19413 MW; 0E07DAB1E23F64A4 CRC64;

Query Match          90.3%; Score 28; DB 2; Length 170;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSRYAR 6
Db 65 FARYAR 70

RESULT 45
ID Q6DRM0_BRARE PRELIMINARY; PRT; 170 AA.
AC Q6DRM0;
DT 16-AUG-2004, integrated into UniProtKB/TrEMBL.
DT 16-AUG-2004, sequence version 1.
DT 07-FEB-2006, entry version 9.
DE Pindall wizard.
DE Name=wrdb;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15256591; DOI=10.1073/pnas.0403929101;
RA Amsterdam A., Nissen R.M., Sun Z., Swindell E.C., Farrington S.,
RA Hopkins N.;
RT "Identification of 315 genes essential for early zebrafish
RT development";
RL Proc. Natl. Acad. Sci. U.S.A. 101:12792-12797(2004).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
DR EMBL: AY648739; AAT68057.1; -; mRNA.
DR Ensembl: ENSDARG0000024641; Danio rerio.
DR ZFIN: ZDB-GENE-030131-7696; wrdb.
DR InterPro: IPR007514; CHDS.
DR Pfam: PF04420; CHDS; 1.
SQ SEQUENCE 170 AA; 19219 MW; 4F4350EB2C29184E CRC64;

Query Match          90.3%; Score 28; DB 2; Length 170;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSRYAR 6
Db 65 FARYAR 70

RESULT 46
ID Q5A4F5_CANAL PRELIMINARY; PRT; 171 AA.
AC Q5A4F5;
DT 26-APR-2005, integrated into UniProtKB/TrEMBL.
DT 26-APR-2005, sequence version 1.
DT 07-FEB-2006, entry version 5.
DE Hypothetical protein.
DE ORFNames=CaO19.7533;
OS Candida albicans SC5314.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=237561;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RA STRAIN=SC5314;
RC
```



```

RX PubMed=15123810; DOI=10.1073/pnas.0401648101;
RA Jones T., Federspiel N.A., Chibana H., Dungan J., Kalman S.,
RA Magee B.B., Newport G., Thorstenson Y.R., Agabidian N., Magee P.T.,
RA Davis R.W., Scherer S.;
RT "The diploid genome sequence of Candida albicans.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:7329-7334(2004).
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL, AAC001000039; EAK9975.1; -; Genomic_DNA.
DR Hypothetical protein.
SQ SEQUENCE 171 AA; 17791 MW; 032B4864D7C8B2E CRC64;

Query Match
Best Local Similarity 83.3%; Score 28; DB 2; Length 171;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRYAR 6
Db 24 FSRYSR 29

RESULT 47
WRB_HUMAN STANDARD; PRT; 174 AA.
AC 000258; O60740;
DT 01-NOV-1997; Integrated into UniProtKB/Swiss-Prot.
DT 01-DEC-2000; sequence version 2.
DT 07-FEB-2006; entry version 34.
DE Tryptophan-rich protein (Congenital heart disease 5 protein).
GN Name=WRB; Synonyms=CHD5;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae;
OC Homo.
NCBI_TaxID=9606;
[1]
RN NUCLEOTIDE SEQUENCE [MRNA], AND SUBCELLULAR LOCATION.
RC TISSUE=Heart;
RX MEDLINE=98204400; PubMed=9544840; DOI=10.1007/s004300050693;
RA Egea A., Mazzocco M., Sotgia F., Arrigo P., Oliva R., Bergonzi S.,
RA Nizetic D., Rasore-Quartino A., Scartezini P.;
RT "Identification and characterization of a new human CDNA from
RT chromosome 21q22.3 encoding a basic nuclear protein.";
RL Hum. Genet. 102:289-293(1998).
[2]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP MEDLINE=20289799; PubMed=10830953; DOI=10.1038/35012518;
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Groner Y.,
RA Soeda E., Ohki M., Takagi T., Sakaki Y., Tautien S., Blechschmidt K.,
RA Polley A., Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA Reichwald K., Rump A., Schillabel M., Schudy A., Zimmermann W.,
RA Rosenblatt A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
RA Minoshima S., Shimizu N., Nordstiek G., Hornischer K., Brandt P.,
RA Schaefer M., Schoen O., Desario A., Reichelt J., Kauer G., Blocker H.,
RA Ramer U., Beck A., Klages S., Hennig S., Rieselmann L., Dagand E.,
RA Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
RA Lehrach H., Reinhardt R., Yaspo M.-L.;
RT "The DNA sequence of human chromosome 21.";
RL Nature 405:311-319(2000).
[3]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RP TISSUE=Ovary;
RX MEDLINE=2338257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

```

```

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.V., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Millaly S.J.,
RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield V.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- SUBCELLULAR LOCATION: Nucleus.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL, Y12478; CAA73081.1; -; mRNA.
DR EMBL, AF064861; -; NOT ANNOTATED CDS; Genomic_DNA.
DR EMBL, AL163279; CAB90454.1; -; Genomic_DNA.
DR EMBL, BC012415; AAH12415.1; -; mRNA.
DR Ensembl, ENSG00000182093; Homo sapiens.
DR H-InvDB, HIX0016117; -.
DR HGNC, HGNC:12790; WRB.
DR MIM, 602915; Gene.
DR GO, GO:0005634; C:nucleus; TAS.
DR InterPro, IPR007514; CHD5.
DR Pfam, PF04420; CHD5; 1.
KW Nuclear protein.
FT CHAIN 1 174
FT
FT CONFLICT 31 31 Tryptophan-rich protein.
FT SEQUENCE 174 AA; 19780 MW; A01F1C564EAB6 CRC64;

Query Match
Best Local Similarity 83.3%; Score 28; DB 1; Length 174;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRYAR 6
Db 69 FARYAR 74

RESULT 48
WRB_MOUSE STANDARD; PRT; 174 AA.
AC 08K0D7; Q9D1W0;
DT 01-MAY-2005; Integrated into UniProtKB/Swiss-Prot.
DT 01-OCT-2002; sequence version 1.
DT 07-FEB-2006; entry version 18.
DE Tryptophan-rich protein.
GN Name=WRB;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 2).
RP STRAIN=C57BL/6J; TISSUE=Corpus striatum;
RX PubMed=16141072; DOI=10.1126/science.11112014;
RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
RA Bajic V.B., Bremner S.E., Batalov S., Forrest A.R., Zavolan M.,
RA Davis M.J., Wilming L.G., Aldins V., Allen J.E.,
RA Ambesi-Impombato A., Apweiler R., Aturaliya R.N., Bailey T.L.,
RA Bansal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,

```



```
OC Muroidea; Muridae; Murinae; Rattus.
OK NCBI_TaxID=10116;
RN (1)
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] .
RC TISSUE=Prostate;
RG NIH - Mammalian Gene Collection (MGC) project;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Nucleus (By similarity).
-----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
-----
DR EMBL: BC062049.1; -; mRNA.
DR Ensembl: ENSRNOG0000001629; Rattus norvegicus.
DR RGD: 735104; WtB.
DR InterPro: IPR007514; CHD5.
DR Pfam: PF04420; CHD5; 1.
KW Nuclear protein.
FT CHAIN 1 174 Tryptophan-rich protein.
      /FTId=PRO_0000065982.
SQ SEQUENCE 174 AA; 19899 MW; 7941B91831F2DC6B CRC64;

Query Match          90.3%; Score 28; DB 1; Length 174;
Best Local Similarity 83.3%; Pred. No. 2.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 FSRYAR 6
   |:|||
Db 69 FARYAR 74
```

Search completed: August 29, 2006, 06:13:22
Job time : 65 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2006, 06:14:29 ; Search time 10.3366 Seconds
(Without alignments)
50.808 Million cell updates/sec

Title: US-10-541-343-2

Perfect score: 31

Sequence: 1 FSRVYR 6

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database :

Issued Patents AA:*

1: /EMC_Celerra_SIDS3/prodata/2/iaa/5/COMB.pep:*

2: /EMC_Celerra_SIDS3/prodata/2/iaa/6/COMB.pep:*

3: /EMC_Celerra_SIDS3/prodata/2/iaa/7/COMB.pep:*

4: /EMC_Celerra_SIDS3/prodata/2/iaa/H/COMB.pep:*

5: /EMC_Celerra_SIDS3/prodata/2/iaa/PC/COMB.pep:*

6: /EMC_Celerra_SIDS3/prodata/2/iaa/RE/COMB.pep:*

7: /EMC_Celerra_SIDS3/prodata/2/iaa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	100.0	33	2	US-08-928-862-4
2	31	100.0	33	2	US-09-592-018-4
3	31	100.0	33	2	US-09-989-481-4
4	31	100.0	357	2	US-09-198-452A-977
5	31	100.0	359	2	US-09-438-185A-906
6	31	100.0	466	2	US-08-914-375C-60
7	31	100.0	509	2	US-09-252-991A-23341
8	28	90.3	188	2	US-09-949-016-11070
9	28	90.3	1778	2	US-09-252-991A-18159
10	27	87.1	198	2	US-09-252-991A-12051
11	27	87.1	384	2	US-09-252-991A-19881
12	27	87.1	387	2	US-09-489-039A-14027
13	27	87.1	456	2	US-08-914-375C-61
14	27	87.1	460	2	US-09-198-452A-538
15	27	87.1	460	2	US-09-489-039A-13505
16	27	87.1	469	2	US-09-489-039A-8224
17	27	87.1	484	3	US-10-026-140-2
18	27	87.1	538	2	US-09-134-001C-4633
19	27	87.1	578	2	US-09-438-185A-599
20	27	87.1	630	2	US-09-602-787A-548
21	27	87.1	650	2	US-09-949-016-10468
22	27	87.1	675	2	US-09-252-991A-29161
23	26	83.9	108	2	US-09-543-681A-5306
24	26	83.9	111	2	US-09-543-681A-6207
25	26	83.9	119	1	US-08-497-312-20
26	26	83.9	119	1	US-08-497-312-22

27	26	83.9	119	2	US-08-875-674A-1	Sequence 1, Appli
28	26	83.9	119	2	US-08-875-674A-3	Sequence 3, Appli
29	26	83.9	130	2	US-10-118-100-51	Sequence 51, Appli
30	26	83.9	131	2	US-09-270-767-44176	Sequence 44176, A
31	26	83.9	165	2	US-09-270-767-36842	Sequence 36842, A
32	26	83.9	165	2	US-09-270-767-52059	Sequence 52059, A
33	26	83.9	229	2	US-09-248-528-11	Sequence 11, Appli
34	26	83.9	229	2	US-09-549-108-11	Sequence 11, Appli
35	26	83.9	229	2	US-09-549-111-11	Sequence 11, Appli
36	26	83.9	229	2	US-09-549-106-11	Sequence 11, Appli
37	26	83.9	229	2	US-09-550-394-11	Sequence 11, Appli
38	26	83.9	233	2	US-09-489-039A-12718	Sequence 12718, A
39	26	83.9	252	2	US-09-949-016-6472	Sequence 6472, Ap
40	26	83.9	253	2	US-09-949-016-9375	Sequence 9375, Ap
41	26	83.9	264	2	US-09-198-452A-265	Sequence 265, App
42	26	83.9	270	2	US-09-438-185A-256	Sequence 256, App
43	26	83.9	276	2	US-09-252-991A-25159	Sequence 25159, A
44	26	83.9	283	2	US-09-252-991A-20485	Sequence 20485, A
45	26	83.9	299	2	US-08-944-483-66	Sequence 66, Appli
46	26	83.9	308	2	US-09-248-796A-14594	Sequence 14594, A
47	26	83.9	319	2	US-09-386-642-12	Sequence 12, Appli
48	26	83.9	327	2	US-09-489-039A-14005	Sequence 14005, A
49	26	83.9	328	2	US-09-386-642-11	Sequence 11, Appli
50	26	83.9	337	2	US-09-438-185A-960	Sequence 960, App
51	26	83.9	340	2	US-09-134-001C-4056	Sequence 4056, Ap
52	26	83.9	341	2	US-09-328-352-7664	Sequence 7664, Ap
53	26	83.9	342	2	US-10-037-417-133	Sequence 133, App
54	26	83.9	343	2	US-10-037-417-134	Sequence 134, App
55	26	83.9	343	2	US-09-948-094-2	Sequence 2, Appli
56	26	83.9	343	2	US-10-037-417-130	Sequence 130, App
57	26	83.9	363	2	US-09-252-991A-32783	Sequence 32783, A
58	26	83.9	376	2	US-09-902-540-11373	Sequence 11373, A
59	26	83.9	393	2	US-09-543-681A-6467	Sequence 6467, Ap
60	26	83.9	435	3	US-10-000-512-18	Sequence 18, Appli
61	26	83.9	451	2	US-09-270-767-40751	Sequence 40751, A
62	26	83.9	451	2	US-09-270-767-55967	Sequence 55967, A
63	26	83.9	475	2	US-09-489-039A-10154	Sequence 10154, A
64	26	83.9	504	2	US-09-583-110-5088	Sequence 5088, Ap
65	26	83.9	504	2	US-09-107-433-3732	Sequence 3732, Ap
66	26	83.9	506	2	US-09-489-039A-8177	Sequence 8177, Ap
67	26	83.9	544	2	US-09-252-991A-28995	Sequence 28995, A
68	26	83.9	547	2	US-09-489-039A-10749	Sequence 10749, A
69	26	83.9	571	2	US-10-036-041-23	Sequence 23, Appli
70	26	83.9	577	2	US-10-104-047-2569	Sequence 2569, Ap
71	26	83.9	644	2	US-09-543-681A-6171	Sequence 6171, Ap
72	26	83.9	675	2	US-09-489-039A-9046	Sequence 9046, Ap
73	26	83.9	769	2	US-09-543-681A-7175	Sequence 7175, Ap
74	26	83.9	945	2	US-09-198-452A-1030	Sequence 1030, Ap
75	26	83.9	1011	2	US-09-538-092-1106	Sequence 1106, Ap
76	26	83.9	1852	2	US-10-023-649A-2	Sequence 2, Appli
77	25	80.6	38	7	US-14590-5	Sequence No. 5614590
78	25	80.6	70	2	US-09-149-476-581	Sequence 581, App
79	25	80.6	109	2	US-09-583-110-3759	Sequence 3759, Ap
80	25	80.6	124	2	US-09-134-000C-4005	Sequence 4005, Ap
81	25	80.6	124	2	US-09-270-767-40656	Sequence 40656, A
82	25	80.6	124	2	US-09-270-767-55872	Sequence 55872, A
83	25	80.6	128	2	US-09-198-452A-795	Sequence 795, App
84	25	80.6	129	2	US-09-107-532A-6571	Sequence 6571, Ap
85	25	80.6	129	2	US-09-902-540-12261	Sequence 12261, Ap
86	25	80.6	132	2	US-09-107-433-3061	Sequence 3061, Ap
87	25	80.6	135	2	US-09-732-210-406	Sequence 406, App
88	25	80.6	138	2	US-09-252-991A-24037	Sequence 24037, A
89	25	80.6	144	2	US-09-270-767-47106	Sequence 47106, A
90	25	80.6	155	2	US-09-252-991A-30401	Sequence 30401, A
91	25	80.6	155	2	US-09-489-039A-10985	Sequence 10985, A
92	25	80.6	156	2	US-09-252-991A-26582	Sequence 26582, A
93	25	80.6	161	2	US-09-252-991A-16868	Sequence 16868, A
94	25	80.6	181	2	US-09-902-540-13247	Sequence 13247, A
95	25	80.6	192	2	US-09-252-991A-19994	Sequence 19904, A
96	25	80.6	200	1	US-08-606-143-42	Sequence 42, Appli
97	25	80.6	213	1	US-08-531-525-12	Sequence 12, Appli
98	25	80.6	213	1	US-08-718-270A-12	Sequence 12, Appli
99	25	80.6	216	2	US-09-252-991A-22665	Sequence 22665, A

100	25	80.6	222	2	US-08-956-182-17	Sequence 17, Appl	173	24	77.4	366	2	US-09-252-991A-24466	Sequence 24466, A
101	25	80.6	234	2	US-09-543-681A-4712	Sequence 4712, Ap	174	24	77.4	367	2	US-09-166-205B-68	Sequence 68, Appl
102	25	80.6	252	2	US-09-270-767-33455	Sequence 33455, A	175	24	77.4	367	2	US-09-806-658-4	Sequence 4, Appl1
103	25	80.6	252	2	US-09-270-767-48672	Sequence 48672, A	176	24	77.4	376	2	US-09-712-363-239	Sequence 239, Ap
104	25	80.6	259	2	US-09-252-991A-31123	Sequence 31123, A	177	24	77.4	381	2	US-09-134-000C-5229	Sequence 5229, Ap
105	25	80.6	270	2	US-09-248-796A-18267	Sequence 18267, A	178	24	77.4	396	2	US-09-902-540-10455	Sequence 14455, A
106	25	80.6	297	7	5514590-2	Patent No. 5514590	179	24	77.4	429	2	US-09-270-767-12603	Sequence 42603, A
107	25	80.6	321	2	US-09-198-452A-401	Sequence 401, App	180	24	77.4	446	2	US-10-034-749-2105	Sequence 2105, Ap
108	25	80.6	321	2	US-09-438-185A-385	Sequence 385, App	181	24	77.4	446	2	US-10-094-749-1068	Sequence 3068, Ap
109	25	80.6	369	2	US-09-543-681A-7514	Sequence 7514, Ap	182	24	77.4	449	2	US-08-914-375C-68	Sequence 68, Appl
110	25	80.6	466	2	US-09-328-352-7117	Sequence 7117, Ap	183	24	77.4	463	2	US-09-489-039A-10537	Sequence 10537, A
111	25	80.6	470	2	US-09-583-110-4689	Sequence 4689, Ap	184	24	77.4	469	2	US-09-489-039A-8689	Sequence 8689, Ap
112	25	80.6	473	2	US-09-107-433-4669	Sequence 4669, Ap	185	24	77.4	477	2	US-09-252-991A-29825	Sequence 29825, A
113	25	80.6	524	2	US-09-252-991A-28370	Sequence 28370, A	186	24	77.4	477	2	US-09-489-039A-9937	Sequence 9937, Ap
114	25	80.6	527	2	US-09-543-681A-6795	Sequence 6795, Ap	187	24	77.4	484	2	US-09-489-039A-9895	Sequence 9895, Ap
115	25	80.6	543	2	US-09-438-185A-7948	Sequence 7948, App	188	24	77.4	492	2	US-09-338-352-6875	Sequence 6875, Ap
116	25	80.6	654	2	US-09-488-039A-7948	Sequence 7948, Ap	189	24	77.4	499	2	US-09-437-568A-42	Sequence 42, Appl
117	25	80.6	658	1	US-08-190-802A-34	Sequence 34, Appl	190	24	77.4	501	2	US-09-270-767-4373	Sequence 4373, A
118	25	80.6	658	2	US-08-477-346-34	Sequence 34, Appl	191	24	77.4	511	2	US-09-543-681A-6450	Sequence 6450, Ap
119	25	80.6	658	2	US-08-473-089-34	Sequence 34, Appl	192	24	77.4	512	2	US-09-716-865-14	Sequence 14, Appl
120	25	80.6	658	2	US-08-487-072A-34	Sequence 34, Appl	193	24	77.4	523	2	US-09-538-092-571	Sequence 571, App
121	25	80.6	822	2	US-09-826-312A-12	Sequence 12, Appl	194	24	77.4	530	2	US-09-270-767-46567	Sequence 46567, A
122	25	80.6	822	2	US-09-542-497A-12	Sequence 12, Appl	195	24	77.4	531	2	US-09-489-039A-9781	Sequence 9781, Ap
123	25	80.6	822	2	US-10-108-767-12	Sequence 12, Appl	196	24	77.4	536	2	US-09-252-991A-20771	Sequence 20771, A
124	25	80.6	822	2	US-10-152-156-12	Sequence 12, Appl	197	24	77.4	560	2	US-09-489-039A-11540	Sequence 11540, A
125	25	80.6	822	3	US-10-835-096-12	Sequence 12, Appl	198	24	77.4	563	2	US-09-437-568A-21	Sequence 21, Appl
126	25	80.6	850	2	US-09-893-525-42	Sequence 42, Appl	199	24	77.4	575	2	US-09-252-991A-26328	Sequence 26328, A
127	25	80.6	921	2	US-09-543-681A-5734	Sequence 5734, Ap	200	24	77.4	587	2	US-09-902-540-15496	Sequence 15496, A
128	25	80.6	925	1	US-08-252-995D-4	Sequence 4, Appl1	201	24	77.4	719	2	US-09-252-991A-22278	Sequence 22278, A
129	25	80.6	925	1	US-08-834-108-4	Sequence 4, Appl1	202	24	77.4	732	2	US-10-070-634-6	Sequence 6, Appl1
130	25	80.6	943	2	US-10-104-047-2552	Sequence 2552, Ap	203	24	77.4	1142	2	US-09-252-991A-18234	Sequence 18234, A
131	25	80.6	1077	2	US-09-949-016-11650	Sequence 11650, A	204	24	77.4	1456	1	US-08-803-973-2	Sequence 2, Appl1
132	25	80.6	1427	2	US-09-538-092-1044	Sequence 1044, Ap	205	24	77.4	1456	1	US-08-803-972-2	Sequence 2, Appl1
133	25	80.6	2628	2	US-09-413-814-11	Sequence 11, Appl	206	24	77.4	1968	1	US-07-745-206A-7	Sequence 7, Appl1
134	25	80.6	2802	2	US-09-542-331-1	Sequence 1, Appl1	207	24	77.4	1968	1	US-08-455-543A-45	Sequence 45, Appl
135	25	80.6	2802	2	US-09-510-791-1	Sequence 1, Appl1	208	24	77.4	1968	1	US-08-223-305C-45	Sequence 45, Appl
136	24	77.4	14	1	US-08-465-325-16	Sequence 126, App	209	24	77.4	1968	1	US-08-311-363-7	Sequence 7, Appl1
137	24	77.4	14	2	US-09-115-737-126	Sequence 126, App	210	24	77.2	10	2	US-08-159-339A-296	Sequence 296, App
138	24	77.4	95	2	US-09-252-991A-25355	Sequence 25355, A	211	23	74.2	11	2	US-08-403-459-17	Sequence 37, Appl
139	24	77.4	101	2	US-09-134-001C-4636	Sequence 4636, Ap	212	23	74.2	13	2	US-08-403-459-40	Sequence 40, Appl
140	24	77.4	107	2	US-09-248-796A-26188	Sequence 26188, A	213	23	74.2	14	1	US-08-467-083-64	Sequence 64, Appl
141	24	77.4	136	2	US-09-489-039A-8896	Sequence 8896, Ap	214	23	74.2	14	1	US-08-414-417B-64	Sequence 64, Appl
142	24	77.4	162	2	US-09-732-210-43	Sequence 43, Appl	215	23	74.2	14	1	US-08-466-348A-64	Sequence 64, Appl
143	24	77.4	186	2	US-09-489-039A-9863	Sequence 9863, Ap	216	23	74.2	14	1	US-08-468-545B-64	Sequence 64, Appl
144	24	77.4	215	2	US-09-252-991A-17696	Sequence 17696, A	217	23	74.2	14	2	US-08-466-680B-64	Sequence 64, Appl
145	24	77.4	249	2	US-09-252-991A-30185	Sequence 30185, A	218	23	74.2	14	2	US-09-027-998A-11	Sequence 11, Appl
146	24	77.4	283	2	US-09-583-110-4252	Sequence 4252, Ap	219	23	74.2	14	2	US-08-403-459-18	Sequence 38, Appl
147	24	77.4	286	2	US-09-107-532A-6850	Sequence 6850, Ap	220	23	74.2	14	2	US-08-403-459-39	Sequence 39, Appl
148	24	77.4	287	2	US-09-710-279-2538	Sequence 2538, Ap	221	23	74.2	14	2	US-08-403-459-65	Sequence 65, Appl
149	24	77.4	291	2	US-09-489-039A-7281	Sequence 7281, Ap	222	23	74.2	14	2	US-08-403-459-66	Sequence 66, Appl
150	24	77.4	302	2	US-09-270-767-59138	Sequence 59138, A	223	23	74.2	14	2	US-09-354-533-64	Sequence 64, Appl
151	24	77.4	307	2	US-09-489-039A-9062	Sequence 9062, Ap	224	23	74.2	15	1	US-09-943-692-11	Sequence 11, Appl
152	24	77.4	324	2	US-09-107-532A-6123	Sequence 6123, Ap	225	23	74.2	15	1	US-08-467-083-60	Sequence 60, Appl
153	24	77.4	331	1	US-08-180-209B-55	Sequence 55, Appl	226	23	74.2	15	1	US-08-414-417B-50	Sequence 50, Appl
154	24	77.4	331	1	US-08-180-209B-57	Sequence 57, Appl	227	23	74.2	15	1	US-08-468-545B-50	Sequence 50, Appl
155	24	77.4	331	1	US-09-008-962-5	Sequence 5, Appl1	228	23	74.2	15	1	US-08-468-545B-50	Sequence 50, Appl
156	24	77.4	331	1	US-08-675-507-5	Sequence 5, Appl1	229	23	74.2	15	2	US-08-466-680B-50	Sequence 50, Appl
157	24	77.4	331	2	US-09-213-205-5	Sequence 5, Appl1	230	23	74.2	15	2	US-09-354-533-50	Sequence 50, Appl
158	24	77.4	331	2	US-08-474-883-55	Sequence 55, Appl	231	23	74.2	16	2	US-09-027-998A-20	Sequence 20, Appl
159	24	77.4	331	2	US-08-166-205B-55	Sequence 55, Appl	232	23	74.2	16	2	US-09-943-692-20	Sequence 20, Appl
160	24	77.4	331	2	US-09-166-205B-57	Sequence 57, Appl	233	23	74.2	17	2	US-08-197-464-11	Sequence 11, Appl
161	24	77.4	331	2	US-09-166-205B-57	Sequence 57, Appl	234	23	74.2	17	5	PCT-US95-02121-11	Sequence 11, Appl
162	24	77.4	331	2	US-09-166-205B-70	Sequence 70, Appl	235	23	74.2	27	1	US-08-776-815B-14	Sequence 14, Appl
163	24	77.4	331	2	US-09-806-658-11	Sequence 11, Appl	236	23	74.2	27	1	US-08-776-815B-15	Sequence 15, Appl
164	24	77.4	331	2	US-09-806-658-12	Sequence 12, Appl	237	23	74.2	27	1	US-08-776-815B-15	Sequence 15, Appl
165	24	77.4	331	2	US-09-902-540-16733	Sequence 16733, A	238	23	74.2	27	2	US-08-737-629-3	Sequence 3, Appl1
166	24	77.4	331	5	PCT-US94-02629-55	Sequence 55, Appl	239	23	74.2	28	1	US-08-776-815B-2	Sequence 2, Appl1
167	24	77.4	331	5	PCT-US94-02629-57	Sequence 57, Appl	240	23	74.2	29	1	US-08-776-815B-3	Sequence 3, Appl1
168	24	77.4	333	2	US-09-583-110-4033	Sequence 4033, Ap	241	23	74.2	29	1	US-08-776-815B-4	Sequence 4, Appl1
169	24	77.4	334	2	US-09-107-433-3718	Sequence 3718, Ap	242	23	74.2	30	1	US-08-776-815B-5	Sequence 5, Appl1
170	24	77.4	350	2	US-09-248-796A-18438	Sequence 18438, A	243	23	74.2	30	1	US-08-776-815B-6	Sequence 6, Appl1
171	24	77.4	355	2	US-09-248-796A-20711	Sequence 20711, A	244	23	74.2	30	2	US-08-776-815B-7	Sequence 7, Appl1
172	24	77.4	360	2	US-09-252-991A-31993	Sequence 31993, A	245	23	74.2	30	2	US-09-270-767-45041	Sequence 45041, A

245	23	74.2	30	2	US-09-270-767-60533	Sequence 60533, A	319	23	74.2	259	2	US-09-270-767-46318	Sequence 46318, A
247	23	74.2	30	2	US-09-623-548A-1300	Sequence 1300, Ap	320	23	74.2	251	7	5320958	Patent No. 5320958
248	23	74.2	30	2	US-09-657-276-1300	Sequence 1300, Ap	321	23	74.2	255	1	US-07-857-224B-66	Sequence 66, Appl
249	23	74.2	39	2	US-09-270-767-58754	Sequence 38754, A	322	23	74.2	257	2	US-09-489-039A-11725	Sequence 11725, A
250	23	74.2	39	2	US-09-270-767-53971	Sequence 53971, A	323	23	74.2	270	2	US-09-724-623-73	Sequence 73, Appl
251	23	74.2	45	1	US-08-056-200-98	Sequence 98, Appl	324	23	74.2	270	2	US-09-634-238-269	Sequence 269, App
252	23	74.2	45	1	US-08-800-644-98	Sequence 98, Appl	325	23	74.2	270	2	US-09-634-238-418	Sequence 418, App
253	23	74.2	47	1	US-08-625-322-22	Sequence 22, Appl	326	23	74.2	270	3	US-10-288-930-73	Sequence 73, Appl
254	23	74.2	47	1	US-08-625-322-23	Sequence 23, Appl	327	23	74.2	265	2	US-09-830-230A-217	Sequence 217, App
255	23	74.2	48	1	US-08-625-322-25	Sequence 25, Appl	328	23	74.2	289	2	US-09-184-658-63	Sequence 63, Appl
256	23	74.2	65	2	US-09-134-001C-2857	Sequence 2857, Ap	329	23	74.2	289	2	US-09-504-262D-63	Sequence 63, Appl
257	23	74.2	67	2	US-09-248-796A-23905	Sequence 23905, A	330	23	74.2	296	2	US-09-515-806-19	Sequence 19, Appl
258	23	74.2	68	2	US-09-270-767-5068	Sequence 35068, A	331	23	74.2	302	2	US-09-599-3608-75	Sequence 75, Appl
259	23	74.2	68	2	US-09-270-767-50285	Sequence 50285, A	332	23	74.2	304	2	US-09-543-681A-8069	Sequence 8069, Ap
260	23	74.2	74	7	5320958-12	Patent No. 5320958	333	23	74.2	308	2	US-09-252-991A-31739	Sequence 31739, A
261	23	74.2	82	1	US-07-881-075-19	Sequence 5490, Ap	334	23	74.2	311	2	US-09-248-796A-20384	Sequence 20384, A
262	23	74.2	82	1	US-08-120-827-19	Sequence 19, Appl	335	23	74.2	312	2	US-09-107-532A-6219	Sequence 6219, Ap
263	23	74.2	82	1	US-08-478-675-19	Sequence 19, Appl	336	23	74.2	314	2	US-09-270-767-38524	Sequence 38524, A
264	23	74.2	82	1	US-08-478-675-19	Sequence 19, Appl	337	23	74.2	314	2	US-09-270-767-53741	Sequence 53741, A
265	23	74.2	92	2	US-09-328-352-5171	Sequence 5171, Ap	338	23	74.2	315	2	US-09-958-969-2	Sequence 2, Appli
266	23	74.2	92	2	US-09-489-039A-13903	Sequence 13903, A	339	23	74.2	320	2	US-09-543-681A-6978	Sequence 6978, Ap
267	23	74.2	92	2	US-09-270-767-58919	Sequence 58919, A	340	23	74.2	327	2	US-09-252-991A-28744	Sequence 28744, A
268	23	74.2	94	1	US-08-467-822-38	Sequence 38, Appl	341	23	74.2	329	2	US-09-489-039A-8390	Sequence 8390, Ap
269	23	74.2	94	2	US-08-432-697-38	Sequence 38, Appl	342	23	74.2	330	2	US-08-778-717-21	Sequence 21, Appl
270	23	74.2	94	2	US-08-466-248-38	Sequence 38, Appl	343	23	74.2	331	2	US-09-710-279-3256	Sequence 3256, Ap
271	23	74.2	95	2	US-09-919-172-102	Sequence 102, App	344	23	74.2	340	2	US-09-540-236-2966	Sequence 2966, Ap
272	23	74.2	95	2	US-09-976-594-667	Sequence 467, App	345	23	74.2	342	1	US-08-454-196-2	Sequence 2, Appli
273	23	74.2	96	2	US-09-902-540-10579	Sequence 10579, A	346	23	74.2	342	2	US-09-064-033-3	Sequence 2, Appli
274	23	74.2	102	2	US-09-949-016-10557	Sequence 10557, A	347	23	74.2	342	2	US-09-291-046-2	Sequence 2, Appli
275	23	74.2	107	2	US-09-513-999C-5491	Sequence 5491, Ap	348	23	74.2	343	1	US-08-454-196-6	Sequence 6, Appli
276	23	74.2	114	2	US-09-252-991A-17880	Sequence 17880, A	349	23	74.2	343	1	US-08-286-819A-4	Sequence 4, Appli
277	23	74.2	116	2	US-09-184-658-48	Sequence 48, Appl	350	23	74.2	343	2	US-08-980-357-4	Sequence 4, Appli
278	23	74.2	116	2	US-09-504-262D-48	Sequence 48, Appl	351	23	74.2	343	2	US-09-064-033-6	Sequence 6, Appli
279	23	74.2	117	2	US-09-902-540-15284	Sequence 15284, A	352	23	74.2	343	2	US-09-291-046-6	Sequence 6, Appli
280	23	74.2	120	1	US-08-497-312-26	Sequence 26, Appl	353	23	74.2	343	2	US-09-357-735-6	Sequence 4, Appli
281	23	74.2	137	2	US-09-902-540-14040	Sequence 14040, A	354	23	74.2	348	2	US-09-107-532A-5305	Sequence 5305, Ap
282	23	74.2	138	2	US-09-270-767-32892	Sequence 32892, A	355	23	74.2	353	2	US-09-328-352-6721	Sequence 6721, Ap
283	23	74.2	138	2	US-09-270-767-48109	Sequence 48109, A	356	23	74.2	358	2	US-10-166-225A-109	Sequence 109, App
284	23	74.2	142	1	US-08-619-598-1	Sequence 1, Appli	357	23	74.2	359	1	US-07-881-075-2	Sequence 2, Appli
285	23	74.2	148	2	US-09-107-532A-3921	Sequence 3921, Ap	358	23	74.2	359	1	US-08-478-675-2	Sequence 2, Appli
286	23	74.2	150	2	US-09-252-991A-24718	Sequence 24718, A	359	23	74.2	360	2	US-09-252-991A-25388	Sequence 25388, A
287	23	74.2	156	2	US-09-387-286-19	Sequence 19, Appl	360	23	74.2	360	2	US-09-710-279-2150	Sequence 2150, Ap
288	23	74.2	160	2	US-09-957-641A-19	Sequence 19, Appl	361	23	74.2	365	2	US-09-949-016-7637	Sequence 7637, Ap
289	23	74.2	161	2	US-09-270-767-43460	Sequence 43460, A	362	23	74.2	365	2	US-09-543-681A-4553	Sequence 4553, Ap
290	23	74.2	162	2	US-09-387-286-18	Sequence 18, Appl	363	23	74.2	366	2	US-09-949-016-9429	Sequence 9429, Ap
291	23	74.2	164	2	US-09-270-767-46444	Sequence 46444, A	364	23	74.2	367	2	US-09-134-001C-3678	Sequence 3678, A
292	23	74.2	167	2	US-10-104-047-3129	Sequence 3129, Ap	365	23	74.2	370	2	US-10-081-644-2	Sequence 2, Appli
293	23	74.2	169	2	US-09-270-767-61958	Sequence 61958, A	366	23	74.2	371	2	US-09-270-767-13350	Sequence 43550, A
294	23	74.2	178	2	US-09-902-540-14484	Sequence 14484, A	367	23	74.2	375	2	US-09-107-532A-3884	Sequence 3884, Ap
295	23	74.2	181	2	US-10-080-960-17	Sequence 17, Appl	368	23	74.2	376	2	US-07-881-075-51	Sequence 51, Appl
296	23	74.2	188	2	US-09-252-991A-17818	Sequence 17818, A	369	23	74.2	380	1	US-07-881-075-51	Sequence 51, Appl
297	23	74.2	190	2	US-09-538-092-954	Sequence 954, App	370	23	74.2	380	1	US-08-120-827-51	Sequence 51, Appl
298	23	74.2	195	2	US-09-270-767-45424	Sequence 45424, A	371	23	74.2	380	1	US-08-478-675-51	Sequence 51, Appl
299	23	74.2	202	2	US-09-252-991A-25553	Sequence 25553, A	372	23	74.2	381	2	US-09-538-092-1017	Sequence 1017, Ap
300	23	74.2	205	2	US-09-724-797-50	Sequence 50, Appl	373	23	74.2	381	2	US-10-104-047-2169	Sequence 2169, Ap
301	23	74.2	210	2	US-09-171-461-34	Sequence 34, Appl	374	23	74.2	384	2	US-09-134-001C-5242	Sequence 5242, Ap
302	23	74.2	215	1	US-09-970-711-34	Sequence 34, Appli	375	23	74.2	385	5	PCT-US93-08528-31	Sequence 31, Appl
303	23	74.2	215	1	US-08-312-870-5	Sequence 5, Appli	376	23	74.2	385	5	PCT-US93-08528-31	Sequence 31, Appl
304	23	74.2	217	2	US-09-902-540-12408	Sequence 12408, A	377	23	74.2	385	5	PCT-US93-08528-31	Sequence 31, Appl
305	23	74.2	221	2	US-09-543-681A-4228	Sequence 4228, Ap	378	23	74.2	395	2	US-09-248-796A-16028	Sequence 16028, A
306	23	74.2	223	2	US-08-778-717-4	Sequence 4, Appli	379	23	74.2	398	1	US-08-630-822A-56	Sequence 56, Appl
307	23	74.2	228	2	US-09-489-039A-11967	Sequence 11967, A	380	23	74.2	398	1	US-09-005-069-56	Sequence 22650, A
308	23	74.2	234	2	US-09-248-796A-20570	Sequence 20570, A	381	23	74.2	400	2	US-09-248-796A-23650	Sequence 23650, A
309	23	74.2	235	2	US-09-538-092-464	Sequence 464, App	382	23	74.2	402	2	US-09-489-039A-10525	Sequence 10525, A
310	23	74.2	243	2	US-09-902-540-15610	Sequence 15610, A	383	23	74.2	402	2	US-09-902-540-16610	Sequence 16610, A
311	23	74.2	244	2	US-09-710-279-994	Sequence 994, App	384	23	74.2	404	2	US-09-252-991A-24433	Sequence 24433, A
312	23	74.2	253	2	US-09-830-230A-218	Sequence 218, App	385	23	74.2	410	2	US-09-543-681A-8105	Sequence 8105, Ap
313	23	74.2	254	2	US-09-247-373B-38	Sequence 38, Appl	386	23	74.2	414	2	US-09-107-433-3915	Sequence 3915, Ap
314	23	74.2	254	2	US-09-949-016-10670	Sequence 10670, A	387	23	74.2	415	2	US-09-554-999-2	Sequence 2, Appli
315	23	74.2	257	2	US-09-489-039A-7196	Sequence 7196, Ap	388	23	74.2	415	2	US-10-188-586A-2	Sequence 2, Appli
316	23	74.2	259	2	US-09-633-947B-3	Sequence 3, Appli	389	23	74.2	416	2	US-09-758-759-145	Sequence 145, App
317	23	74.2	259	2	US-09-543-681A-7362	Sequence 7362, Ap	390	23	74.2	422	2	US-09-949-016-9757	Sequence 9757, A
318	23	74.2	259	2			391	23	74.2	423	2	US-09-270-767-46381	Sequence 46381, A

392	23	74.2	433	1	US-07-661-610C-2	Sequence 2, Appli
393	23	74.2	433	2	US-09-252-991A-30647	Sequence 30647, A
394	23	74.2	434	2	US-09-252-991A-24980	Sequence 24980, A
395	23	74.2	435	2	US-09-902-540-1518	Sequence 1518, A
396	23	74.2	436	2	US-09-252-991A-33064	Sequence 33064, A
397	23	74.2	444	2	US-09-252-991A-31017	Sequence 31017, A
398	23	74.2	446	1	US-08-307-444A-5	Sequence 5, Appli
399	23	74.2	446	1	US-08-587-389-5	Sequence 5, Appli
400	23	74.2	447	2	US-09-252-991A-28081	Sequence 28081, A
401	23	74.2	451	2	US-09-543-681A-4841	Sequence 4841, Ap
402	23	74.2	451	2	US-09-248-796A-20754	Sequence 20754, A
403	23	74.2	453	1	US-08-313-553-11	Sequence 11, Appli
404	23	74.2	453	2	US-08-767-993-11	Sequence 11, Appli
405	23	74.2	456	1	US-08-307-444A-3	Sequence 3, Appli
406	23	74.2	456	1	US-08-307-444A-4	Sequence 4, Appli
407	23	74.2	456	1	US-08-587-389-3	Sequence 3, Appli
408	23	74.2	456	1	US-08-587-389-4	Sequence 4, Appli
409	23	74.2	459	2	US-09-526-309B-1	Sequence 1, Appli
410	23	74.2	459	2	US-09-526-309B-3	Sequence 3, Appli
411	23	74.2	459	2	US-09-526-309B-5	Sequence 5, Appli
412	23	74.2	459	2	US-09-526-309B-7	Sequence 7, Appli
413	23	74.2	459	2	US-09-526-309B-9	Sequence 9, Appli
414	23	74.2	459	2	US-09-526-309B-11	Sequence 11, Appli
415	23	74.2	459	2	US-09-526-309B-13	Sequence 13, Appli
416	23	74.2	459	2	US-09-526-309B-17	Sequence 17, Appli
417	23	74.2	459	2	US-09-526-309B-19	Sequence 19, Appli
418	23	74.2	459	2	US-09-526-309B-21	Sequence 21, Appli
419	23	74.2	459	2	US-09-526-309B-23	Sequence 23, Appli
420	23	74.2	459	2	US-09-526-309B-25	Sequence 25, Appli
421	23	74.2	459	2	US-09-526-309B-27	Sequence 27, Appli
422	23	74.2	459	2	US-09-526-309B-29	Sequence 29, Appli
423	23	74.2	459	2	US-09-526-309B-31	Sequence 31, Appli
424	23	74.2	459	2	US-09-526-309B-33	Sequence 33, Appli
425	23	74.2	459	2	US-09-526-309B-35	Sequence 35, Appli
426	23	74.2	459	2	US-09-526-309B-37	Sequence 37, Appli
427	23	74.2	459	2	US-09-526-309B-39	Sequence 39, Appli
428	23	74.2	459	2	US-09-526-309B-41	Sequence 41, Appli
429	23	74.2	459	2	US-09-526-309B-43	Sequence 43, Appli
430	23	74.2	459	2	US-10-280-858A-1	Sequence 1, Appli
431	23	74.2	459	2	US-10-280-858A-3	Sequence 3, Appli
432	23	74.2	459	2	US-10-280-858A-5	Sequence 5, Appli
433	23	74.2	459	2	US-10-280-858A-7	Sequence 7, Appli
434	23	74.2	459	2	US-10-280-858A-9	Sequence 9, Appli
435	23	74.2	459	2	US-10-280-858A-11	Sequence 11, Appli
436	23	74.2	459	2	US-10-280-858A-13	Sequence 13, Appli
437	23	74.2	459	2	US-10-280-858A-17	Sequence 17, Appli
438	23	74.2	459	2	US-10-280-858A-19	Sequence 19, Appli
439	23	74.2	459	2	US-10-280-858A-21	Sequence 21, Appli
440	23	74.2	459	2	US-10-280-858A-23	Sequence 23, Appli
441	23	74.2	459	2	US-10-280-858A-25	Sequence 25, Appli
442	23	74.2	459	2	US-10-280-858A-27	Sequence 27, Appli
443	23	74.2	459	2	US-10-280-858A-29	Sequence 29, Appli
444	23	74.2	459	2	US-10-280-858A-31	Sequence 31, Appli
445	23	74.2	459	2	US-10-280-858A-33	Sequence 33, Appli
446	23	74.2	459	2	US-10-280-858A-35	Sequence 35, Appli
447	23	74.2	459	2	US-10-280-858A-37	Sequence 37, Appli
448	23	74.2	459	2	US-10-280-858A-39	Sequence 39, Appli
449	23	74.2	459	2	US-10-280-858A-41	Sequence 41, Appli
450	23	74.2	459	2	US-10-280-858A-43	Sequence 43, Appli
451	23	74.2	460	1	US-07-817-920-4	Sequence 4, Appli
452	23	74.2	460	1	US-07-996-772A-9	Sequence 9, Appli
453	23	74.2	460	1	US-08-194-338-13	Sequence 13, Appli
454	23	74.2	460	1	US-08-370-542-4	Sequence 4, Appli
455	23	74.2	460	1	US-08-117-006-4	Sequence 4, Appli
456	23	74.2	460	1	US-08-216-594-4	Sequence 4, Appli
457	23	74.2	460	1	US-08-542-358-4	Sequence 4, Appli
458	23	74.2	460	1	US-09-018-351-4	Sequence 4, Appli
459	23	74.2	460	2	US-09-032-742-4	Sequence 4, Appli
460	23	74.2	460	2	US-09-032-742-20	Sequence 20, Appli
461	23	74.2	460	2	US-09-032-742-23	Sequence 23, Appli
462	23	74.2	460	2	US-09-328-314-18	Sequence 18, Appli
463	23	74.2	460	2	US-09-489-039A-12680	Sequence 12680, A
464	23	74.2	460	5	PCT-US93-00149-4	Sequence 4, Appli

465

23

74.2

461

2

US-09-526-309B-15

Sequence 15, Appli

466

23

74.2

461

2

US-10-280-858A-15

Sequence 15, Appli

467

23

74.2

464

2

US-09-902-540-1518

Sequence 1518, A

468

23

74.2

475

1

US-08-307-444A-2

Sequence 2, Appli

469

23

74.2

475

1

US-08-307-444A-2

Sequence 2, Appli

470

23

74.2

475

1

US-08-587-389-1

Sequence 1, Appli

471

23

74.2

475

1

US-08-587-389-2

Sequence 2, Appli

472

23

74.2

476

1

US-08-014-723-1

Sequence 1, Appli

473

23

74.2

476

1

US-08-014-723-2

Sequence 2, Appli

474

23

74.2

476

1

US-08-014-723-18

Sequence 18, Appli

475

23

74.2

476

1

US-08-110-011A-1

Sequence 1, Appli

476

23

74.2

476

1

US-08-110-011A-2

Sequence 2, Appli

477

23

74.2

476

1

US-08-110-011A-18

Sequence 18, Appli

478

23

74.2

481

2

US-09-252-991A-28020

Sequence 28020, A

479

23

74.2

482

2

US-09-252-991A-20129

Sequence 20129, A

480

23

74.2

485

2

US-09-949-016-10440

Sequence 10440, A

481

23

74.2

485

2

US-10-094-749-3144

Sequence 3144, Ap

482

23

74.2

488

2

US-09-489-039A-9557

Sequence 9557, Ap

483

23

74.2

492

1

US-09-134-000C-3694

Sequence 3694, Ap

484

23

74.2

494

1

US-08-014-723-14

Sequence 14, Appli

485

23

74.2

494

1

US-08-014-723-16

Sequence 16, Appli

486

23

74.2

494

1

US-08-110-011A-14

Sequence 14, Appli

487

23

74.2

494

1

US-08-110-011A-16

Sequence 16, Appli

488

23

74.2

497

1

US-08-312-870-3

Sequence 3, Appli

489

23

74.2

497

2

US-09-331-793-4

Sequence 4, Appli

490

23

74.2

498

1

US-08-733-564-2

Sequence 2, Appli

491

23

74.2

503

2

US-09-562-737-66

Sequence 66, Appli

492

23

74.2

506

2

US-09-489-039A-7307

Sequence 7307, Ap

493

23

74.2

516

2

US-09-509-994-1

Sequence 1, Appli

494

23

74.2

516

2

US-09-509-994-2

Sequence 2, Appli

495

23

74.2

536

2

US-09-107-532A-6943

Sequence 6943, Ap

496

23

74.2

536

2

US-09-653-274-10

Sequence 10, Appli

497

23

74.2

541

1

US-10-461-791-19

Sequence 19, Appli

498

23

74.2

546

1

US-08-484-438-6

Sequence 6, Appli

499

23

74.2

542

2

US-09-302-769-14

Sequence 44, Appli

500

23

74.2

542

2

US-09-712-363-180

Sequence 180, Appli

RESULT 1

US-08-928-862-4

; Sequence 4, Application US/08928862

; Patent No. 6309877

; GENERAL INFORMATION:

; APPLICANT: Chau, Raymond M. W.

; TITLE OF INVENTION: Isolation and Use of Motoneuronotrophic Factors

; FILE REFERENCE: 12592-2

; CURRENT APPLICATION NUMBER: US/08/928, 862

; CURRENT FILING DATE: 1997-09-12

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 4

; LENGTH: 33

; TYPE: PRT

; ORGANISM: Homo sapiens

US-08-928-862-4

Query Match

Best Local Similarity 100.0%; Pred. No. 5.7;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy

1

FSRYAR 6

|||||

Db

17

FSRYAR 22

RESULT 2

US-09-592-018-4

; Sequence 4, Application US/09592018

; Patent No. 6759389

; GENERAL INFORMATION:

APPLICANT: Chau, Raymond M.W.
TITLE OF INVENTION: Isolation and Use of Motoneuronotropic Factors
FILE REFERENCE: 12592-3
CURRENT APPLICATION NUMBER: US/09/592,018
CURRENT FILING DATE: 2000-06-12
PRIOR APPLICATION NUMBER: US 08/928862
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: US 08/751225
PRIOR FILING DATE: 1996-11-15
PRIOR APPLICATION NUMBER: US 60/026792
PRIOR FILING DATE: 1996-09-27
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn version 3.0
SEQ ID NO 4
LENGTH: 33
TYPE: PRT
ORGANISM: Homo Sapiens
US-09-592-018-4

Query Match 100.0%; Score 31; DB 2; Length 33;
Best Local Similarity 100.0%; Pred. No. 5.7;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 FSRVAR 6
Db 17 FSRVAR 22

RESULT 3
US-09-989-481-4
Sequence 4, Application US/09989481
Patent No. 6841531
GENERAL INFORMATION:
APPLICANT: Chau, Raymond M.W.
TITLE OF INVENTION: Isolation and Use of Motoneuronotropic Factors
FILE REFERENCE: 12592-4
CURRENT APPLICATION NUMBER: US/09/989,481
CURRENT FILING DATE: 2001-11-20
PRIOR APPLICATION NUMBER: US 09/633,447
PRIOR FILING DATE: 2000-08-07
PRIOR APPLICATION NUMBER: US 08/9928862
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: US 08/751225
PRIOR FILING DATE: 1996-11-15
PRIOR APPLICATION NUMBER: US 60/026792
PRIOR FILING DATE: 1996-09-27
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn version 3.0
SEQ ID NO 4
LENGTH: 33
TYPE: PRT
ORGANISM: Homo Sapiens
US-09-989-481-4

Query Match 100.0%; Score 31; DB 2; Length 33;
Best Local Similarity 100.0%; Pred. No. 5.7;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 FSRVAR 6
Db 17 FSRVAR 22

RESULT 4
US-09-198-452A-977
Sequence 977, Application US/09198452A
Patent No. 6559294
GENERAL INFORMATION:
APPLICANT: Grifflais, R.
TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
TITLE OF INVENTION: and treatment of infection
FILE REFERENCE: 9710-003-999

CURRENT APPLICATION NUMBER: US/09/198,452A
CURRENT FILING DATE: 1998-11-24
NUMBER OF SEQ ID NOS: 6849
SEQ ID NO 977
LENGTH: 357
TYPE: PRT
ORGANISM: Chlamydia pneumoniae
US-09-198-452A-977

Query Match 100.0%; Score 31; DB 2; Length 357;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 FSRVAR 6
Db 135 FSRVAR 140

RESULT 5
US-09-438-185A-906
Sequence 906, Application US/09438185A
Patent No. 682071
GENERAL INFORMATION:
APPLICANT: Stephens, Richard
APPLICANT: Mitchell, Wayne
APPLICANT: Kalman, Sue
APPLICANT: Davis, Ronald
APPLICANT: The Regents of the University of California
TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence
FILE REFERENCE: 018941-000411US
CURRENT APPLICATION NUMBER: US/09/438,185A
CURRENT FILING DATE: 2002-03-13
PRIOR APPLICATION NUMBER: US 60/108,279
PRIOR FILING DATE: 1998-11-12
PRIOR APPLICATION NUMBER: US 60/128,606
PRIOR FILING DATE: 1999-04-08
NUMBER OF SEQ ID NOS: 1074
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 906
LENGTH: 359
TYPE: PRT
ORGANISM: Chlamydia pneumoniae
FEATURE:
OTHER INFORMATION: CPN0904
US-09-438-185A-906

Query Match 100.0%; Score 31; DB 2; Length 359;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 FSRVAR 6
Db 137 FSRVAR 142

RESULT 6
US-08-914-375C-60
Sequence 60, Application US/08914375C
Patent No. 6377893
GENERAL INFORMATION:
APPLICANT: Steven A. Benner
NUMBER OF SEQUENCES: 74
CORRESPONDENCE ADDRESS:
ADDRESS: Steven A. Benner
STREET: 1501 NW 68th Terrace
CITY: Gainesville
STATE: FL
COUNTRY: United States
ZIP: 32605-4147
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch diskette
COMPUTER: Apple Macintosh

```
OPERATING SYSTEM: Macintosh 7.0
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/914,375C
FILING DATE: 19-AUG-1997
CLASSIFICATION: 702/20
TELECOMMUNICATION INFORMATION:
TELEPHONE: 352 392 7773
TELEFAX: 352 331 0462
INFORMATION FOR SEQ ID NO: 60:
SEQUENCE CHARACTERISTICS:
LENGTH: 466
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: amino acid
ORIGINAL SOURCE:
ORGANISM: Escherichia coli
FEATURE:
OTHER INFORMATION: ascb.ecoli 6-phospho-strand-glucosidase (E.C. 3.2.1.86
SEQUENCE DESCRIPTION: SEQ ID NO: 60:
US-08-914-375C-60
```

```
Query Match          100.0%; Score 31; DB 2; Length 466;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 FSRVAR 6
      |||||
Db      154 FSRVAR 159
```

```
RESULT 7
US-09-252-991A-23341
; Sequence 23341, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23341
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23341
```

```
Query Match          100.0%; Score 31; DB 2; Length 509;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 FSRVAR 6
      |||||
Db      270 FSRVAR 275
```

```
RESULT 8
US-09-949-016-11070
; Sequence 11070, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
```

```
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11070
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11070
```

```
Query Match          90.3%; Score 28; DB 2; Length 188;
Best Local Similarity 83.3%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 FSRVAR 6
      |||||
Db      83 FSRVAR 88
```

```
RESULT 9
US-09-252-991A-18159
; Sequence 18159, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18159
; LENGTH: 1778
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18159
```

```
Query Match          90.3%; Score 28; DB 2; Length 1778;
Best Local Similarity 83.3%; Pred. No. 1.4e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 FSRVAR 6
      |||||
Db      1131 FSRVAR 1136
```

```
RESULT 10
US-09-252-991A-32051
; Sequence 32051, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32051
; LENGTH: 198
; TYPE: PRT
```

ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32051

Query Match 87.1%; Score 27; DB 2; Length 198;
Best Local Similarity 83.3%; Pred. No. 2.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRYAR 6
| | | | |
Db 94 FSRPAR 99

RESULT 11
US-09-252-991A-19881
; Sequence 19881, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252.991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 19881
LENGTH: 364
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19881

Query Match 87.1%; Score 27; DB 2; Length 364;
Best Local Similarity 83.3%; Pred. No. 4.5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FSRYAR 6
| | | | |
Db 351 FORYAR 356

RESULT 12
US-09-489-039A-14027
; Sequence 14027, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489.039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 14027
LENGTH: 387
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-14027

Query Match 87.1%; Score 27; DB 2; Length 387;
Best Local Similarity 83.3%; Pred. No. 4.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRYAR 6
| | | | |
Db 111 FSRYAQ 116

RESULT 13

US-08-914-375C-61
; Sequence 61, Application US/08914375C
; Patent No. 6377893
; GENERAL INFORMATION:

APPLICANT: Steven A. Benner
NUMBER OF SEQUENCES: 74
CORRESPONDENCE ADDRESS:
ADDRESSEE: Steven A. Benner
STREET: 1501 NW 68th Terrace
CITY: Gainesville
STATE: FL
COUNTRY: United States
ZIP: 32605-4147

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch diskette
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.0
SOFTWARE: Microsoft Word

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/914,375C
FILING DATE: 19-Aug-1997
CLASSIFICATION: 702/20
TELECOMMUNICATION INFORMATION:
TELEPHONE: 352 392 7773
TELEFAX: 352 331 0462

INFORMATION FOR SEQ ID NO: 61:
SEQUENCE CHARACTERISTICS:
LENGTH: 456
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: amino acid
ORIGINAL SOURCE:
ORGANISM: Ertwinia chrysanthemi

FEATURE:
OTHER INFORMATION: arbb erwch 6-phospho-strand-glucosidase (E.C. 3.2.1.86
SEQUENCE DESCRIPTION: SEQ ID NO: 61:
US-08-914-375C-61

Query Match 87.1%; Score 27; DB 2; Length 456;
Best Local Similarity 83.3%; Pred. No. 5.7e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FSRYAR 6
| | | | |
Db 150 FERYAR 155

RESULT 14
US-09-198-452A-638
; Sequence 638, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
APPLICANT: Grifflais, R.
TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, preve
FILE REFERENCE: 9710-003-999
CURRENT APPLICATION NUMBER: US/09/198,452A
CURRENT FILING DATE: 1998-11-24
NUMBER OF SEQ ID NOS: 6849
SEQ ID NO 638
LENGTH: 460
TYPE: PRT
ORGANISM: Chlamydia pneumoniae
US-09-198-452A-638

Query Match 87.1%; Score 27; DB 2; Length 460;
Best Local Similarity 83.3%; Pred. No. 5.7e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FSRYAR 6
| | | | |

```

Db          336 FSRVYR 341

RESULT 15
US-09-489-039A-13505
; Sequence 13505, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Bretton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT FILING DATE: US/09/489,039A
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 13505
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13505

Query Match      87.1%; Score 27; DB 2; Length 460;
Best Local Similarity 83.3%; Pred. No. 5.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy          1 FSRVYR 6
Db          78 FSRVYR 83

RESULT 16
US-09-489-039A-8224
; Sequence 8224, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Bretton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT FILING DATE: US/09/489,039A
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 8224
; LENGTH: 469
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8224

Query Match      87.1%; Score 27; DB 2; Length 469;
Best Local Similarity 83.3%; Pred. No. 5.9e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy          1 FSRVYR 6
Db          154 FSRVYR 159

RESULT 17
US-10-026-140-2
; Sequence 2, Application US/10026140
; Patent No. 7005289
; GENERAL INFORMATION:
; APPLICANT: Dunn-Coleman, Nigel
; APPLICANT: Goedegebuur, Frits
; APPLICANT: Ward, Michael
; APPLICANT: Yao, Jian
; TITLE OF INVENTION: BGL5 Beta-Glucosidase and Nucleic Acids
; TITLE OF INVENTION: Encoding the Same
; FILE REFERENCE: GC697

```

```

; CURRENT APPLICATION NUMBER: US/10/026,140
; CURRENT FILING DATE: 2002-03-26
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 484
; TYPE: PRT
; ORGANISM: Trichoderma reesei
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...((484)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-026-140-2

Query Match      87.1%; Score 27; DB 3; Length 484;
Best Local Similarity 83.3%; Pred. No. 6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy          1 FSRVYR 6
Db          148 FSRVYR 153

RESULT 18
US-09-134-001C-4633
; Sequence 4633, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4633
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4633

Query Match      87.1%; Score 27; DB 2; Length 538;
Best Local Similarity 83.3%; Pred. No. 6.7e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy          1 FSRVYR 6
Db          77 FSRVYR 82

RESULT 19
US-09-438-185A-599
; Sequence 599, Application US/09438185A
; Patent No. 6822071
; GENERAL INFORMATION:
; APPLICANT: Stephens, Richard
; APPLICANT: Mitchell, Wayne
; APPLICANT: Kaiman, Sue
; APPLICANT: Davis, Ronald
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence
; FILE REFERENCE: 018941-000411US
; CURRENT APPLICATION NUMBER: US/09/438,185A
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/108,279
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 60/128,606
; PRIOR FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 1074

```

```
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 599
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; OTHER INFORMATION: CPN0597
US-09-438-185A-599

Query Match      87.1%; Score 27; DB 2; Length 579;
Best Local Similarity 83.3%; Pred. No. 7.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 FSRYAR 6      |||||
Db      455 FSRYVR 460

RESULT 20
US-09-602-787A-548
; Sequence 548, Application US/09602787A
; Patent No. 6696561
; GENERAL INFORMATION:
; APPLICANT: Pompeius, Mark
; APPLICANT: Krüger, Burkhard
; APPLICANT: Schüder, Hartwig
; APPLICANT: Zeldner, Oskar
; APPLICANT: Haberhauser, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
; TITLE OF INVENTION: TRANSPORT
; FILE REFERENCE: BGI-125CP
; CURRENT APPLICATION NUMBER: US/09/602,787A
; CURRENT FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: USN 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931454.3
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931478.0
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931563.9
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932122.1
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932124.8
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932125.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932180.9
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932182.5
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932190.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932191.4
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932209.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932212.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932227.9
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932228.7
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932229.5
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932230.9
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932927.3
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933005.0
```

```
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933006.9
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19940764.9
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940765.7
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940766.5
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940830.0
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940831.9
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940832.7
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940833.5
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19941378.9
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941379.7
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941395.9
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19942077.7
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: DE 19942078.5
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: DE 19942079.3
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: DE 19942088.2
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 678
; SEQ ID NO 548
; LENGTH: 630
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-602-787A-548

Query Match      87.1%; Score 27; DB 2; Length 630;
Best Local Similarity 83.3%; Pred. No. 7.9e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 FSRYAR 6      |||||
Db      105 FSRYGR 110

RESULT 21
US-09-949-016-10468
; Sequence 10468, Application US/09949016
; Patent No. 6612339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10468
; LENGTH: 650
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10468

Query Match      87.1%; Score 27; DB 2; Length 650;
```

Best Local Similarity 83.3%; Pred. No. 8.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 FSRVAR 6
Db 283 FSOYAR 288

RESULT 22

US-09-252-991A-29161
; Sequence 29161, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29161
; LENGTH: 675
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29161

Query Match 87.1%; Score 27; DB 2; Length 675;
Best Local Similarity 83.3%; Pred. No. 8.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSRVAR 6
Db 670 FSRFAR 675

RESULT 23

US-09-543-681A-6306
; Sequence 6306, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 6306
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-6306

Query Match 83.9%; Score 26; DB 2; Length 108;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FSRVAR 6
Db 46 FSRYER 51

RESULT 24

US-09-543-681A-6207
; Sequence 6207, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:

; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 6207
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-6207

Query Match 83.9%; Score 26; DB 2; Length 111;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 25

US-08-497-312-20
; Sequence 20, Application US/08497312
; Patent No. 5712120
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Method for obtaining modified
; TITLE OF INVENTION: Immunoglobulins with reduced immunogenicity of murine
; TITLE OF INVENTION: antibody variable domains, compositions containing them.
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CENTRO DE INMUNOLOGIA MOLECULAR
; STREET: 215 Y 15, ATRABAY PLAYA
; CITY: HAVANA
; STATE:
; COUNTRY: CUBA
; ZIP: 11600
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; FILING DATE: 30-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CU 80/94
; FILING DATE: 30-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BOND, LAURENCE B.
; REGISTRATION NUMBER: 30,549
; REFERENCE/DOCKET NUMBER: 262905
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 801/532-1922
; TELEFAX: 801/531-9168
; TELEX: 388961 1PMO4UT
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 119 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
US-08-497-312-20

Query Match 83.9%; Score 26; DB 1; Length 119;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FSRYA 5
|||
Db 29 FSRYA 33

RESULT 26

US-08-497-312-22
; Sequence 22, Application US/08497312
; Patent No. 5712120
; GENERAL INFORMATION:

APPLICANT:
; TITLE OF INVENTION: Method for obtaining modified
; TITLE OF INVENTION: immunoglobulins with reduced immunogenicity of murine
; TITLE OF INVENTION: antibody variable domains, compositions containing them.
; NUMBER OF SEQUENCES: 31

CORRESPONDENCE ADDRESS:

ADDRESSEE: CENTRO DE INMUNOLOGIA MOLECULAR
STREET: 215 Y 15, ATABEX PLAYA
CITY: HAVANA
STATE:
COUNTRY: CUBA

ZIP: 11600

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/497,312

FILING DATE: 30-JUN-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: CU 80/94

FILING DATE: 30-JUN-1994

ATTORNEY/AGENT INFORMATION:

NAME: BOND, LAURENCE B.

REGISTRATION NUMBER: 30,549

REFERENCE/DOCKET NUMBER: 2629US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 801/532-1922

TELEFAX: 801/531-9168

TELEX: 388961 1PM04UT

INFORMATION FOR SEQ ID NO: 22:

SEQUENCE CHARACTERISTICS:

LENGTH: 119 amino acids

TYPE: amino acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: protein

HYPOTHETICAL: NO

US-08-497-312-22

Query Match 83.9%; Score 26; DB 1; Length 119;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FSRYA 5
|||
Db 29 FSRYA 33

RESULT 27

US-08-875-674A-1

; Sequence 1, Application US/08875674A

; Patent No. 6572857

; GENERAL INFORMATION:

APPLICANT: MONTERO CASIMIRO, J. E.

APPLICANT: LOMBARDO VALADARES, J.

APPLICANT: P REZ RODR GUEZ, R.

APPLICANT: SIERRA BL ZQUEZ, P.

APPLICANT: TORO BRAVO, B. R.

; TITLE OF INVENTION: Anti-CD6 monoclonal antibodies and their uses
; NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lackenbach Siegel Marzullo Aronson & Greenspan, P.C.

STREET: One Chase Road

CITY: Scarsdale

STATE: New York

COUNTRY: U.S.A.

ZIP: 10583

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk 3.5" (1.4 MB)

COMPUTER: Compatible PC IBM (80486, 8 M Ram).

OPERATING SYSTEM: Windows 95.

SOFTWARE: Word Perfect 5.0 for Windows 95.

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/875,674A

FILING DATE: 17-July-1997

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/CU96/00004

FILING DATE: 18-NOV-1996

ATTORNEY/AGENT INFORMATION:

NAME: HENRY A. MARZULLO, JR.

REGISTRATION NUMBER: 20,910

REFERENCE/DOCKET NUMBER: P-12

TELECOMMUNICATION INFORMATION:

TELEPHONE: (914) 723-4300

TELEFAX: (914) 723-4301

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 119 Amino acid residues.

TYPE: Amino acid.

STRANDEDNESS: Unknown.

TOPOLOGY: Unknown.

MOLECULE TYPE: Protein

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: -N Terminal fragment.

ORIGINAL SOURCE:

ORGANISM: Mice Balb/C

INDIVIDUAL ISOLATE: for t1A

TISSUE TYPE: Murine hybridoma

IMMEDIATE SOURCE:

CLONE: Sub-clone for t1A

FEATURE:

IDENTIFICATION METHOD: Experimental.

OTHER INFORMATION: Sequence corresponding to the variable region

; Patent No. 6572857

; OTHER INFORMATION: of the heavy chain of the monoclonal antibody recognizing hum.

; US-08-875-674A-1

Query Match 83.9%; Score 26; DB 2; Length 119;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FSRYA 5
|||
Db 29 FSRYA 33

RESULT 28

US-08-875-674A-3

; Sequence 3, Application US/08875674A

; Patent No. 6572857

; GENERAL INFORMATION:

APPLICANT: MONTERO CASIMIRO, J. E.

APPLICANT: LOMBARDO VALADARES, J.

APPLICANT: P REZ RODR GUEZ, R.

APPLICANT: SIERRA BL ZQUEZ, P.

APPLICANT: TORO BRAVO, B. R.

; TITLE OF INVENTION: Anti-CD6 monoclonal antibodies and their uses

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lackenbach Siegel Marzullo Aronson & Greenspan, P.C.

```

; STREET: One Chase Road
; CITY: Scarsdale
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10583
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk 3.5" (1.4 MB)
; COMPUTER: Compatible PC IBM (80486, 8 M Ram)
; OPERATING SYSTEM: Windows 95
; SOFTWARE: Word Perfect 5.0 for Windows 95
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/875,674A
; FILING DATE: 17-July-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/CU96/00004
; FILING DATE: 18-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: HENRY A. MARZULLO, JR.
; REGISTRATION NUMBER: 20,910
; REFERENCE/DOCKET NUMBER: P-12
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (914) 723-4300
; TELEFAX: (914)-723-4301
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 119 Amino acid residues.
; TYPE: Amino acid.
; STRANDEDNESS: Unknown.
; TOPOLOGY: Unknown.
; MOLECULE TYPE: Protein
; HYPOTHEICAL: No
; ANTI-SENSE: No
; FRAGMENT TYPE: -N Terminal fragment.
; IMMEDIATE SOURCE:
; CELL LINE: NSO " SP 2/0 " CHO
; CLONE: Sub-clone for t1A
; FEATURE:
; IDENTIFICATION METHOD: By similarity with known sequence.
; OTHER INFORMATION: Sequence corresponding to the humanized
; Patent No. 6572857
; OTHER INFORMATION: variant of sub-clone for t1A recognizing human CD6, particula
; OTHER INFORMATION: to the variable region of its heavy chain.
; US-08-875-674A-3
;
; Query Match      83.9%; Score 26; DB 2; Length 119;
; Best Local Similarity 100.0%; Pred. No. 2.4e+02;
; Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY      1 FSRVA 5
;         |||||
; Db      29 FSRVA 33
;
; RESULT 29
; US-10-118-100-51
; Sequence 51, Application US/10118100
; Patent No. 6969250
; GENERAL INFORMATION:
; APPLICANT: Bioinvent International AB
; APPLICANT: Soderlind, Ulf
; APPLICANT: Borreback, Carl
; TITLE OF INVENTION: A Method For In Vitro Molecular
; FILE REFERENCE: Mewburn 341711
; CURRENT APPLICATION NUMBER: US/10/118,100
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: US/09/341,711
; PRIOR FILING DATE: 1999-09-21
; PRIOR APPLICATION NUMBER: PCT/GB98/00219
; PRIOR FILING DATE: 1998-01-26

```

```

; PRIOR APPLICATION NUMBER: GB9701425.2
; PRIOR FILING DATE: 1997-01-24
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 51
; LENGTH: 130
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: scfv-B11 Antibody Fragment
; US-10-118-100-51
;
; Query Match      83.9%; Score 26; DB 2; Length 130;
; Best Local Similarity 100.0%; Pred. No. 2.6e+02;
; Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY      1 FSRVA 5
;         |||||
; Db      37 FSRVA 41
;
; RESULT 30
; US-09-270-767-44176
; Sequence 44176, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 44176
; LENGTH: 131
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
; US-09-270-767-44176
;
; Query Match      83.9%; Score 26; DB 2; Length 131;
; Best Local Similarity 100.0%; Pred. No. 2.6e+02;
; Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY      1 FSRVA 5
;         |||||
; Db      96 FSRVA 100
;
; RESULT 31
; US-09-270-767-36842
; Sequence 36842, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 36842
; LENGTH: 165
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
; US-09-270-767-36842
;
; Query Match      83.9%; Score 26; DB 2; Length 165;
; Best Local Similarity 100.0%; Pred. No. 3.3e+02;
; Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```


Qy 1 FSRYA 5
Db 55 FSRYA 59

RESULT 32

US-09-270-767-52059
; Sequence 52059, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-034
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO: 52059
; LENGTH: 165
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-52059

Query Match 83.9%; Score 26; DB 2; Length 165;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSRYA 5
Db 55 FSRYA 59

RESULT 33

US-09-248-528-11
; Sequence 11, Application US/09248528C
; Patent No. 6153415
; GENERAL INFORMATION:
; APPLICANT: Oriol, Patrick J
; APPLICANT: Padmakumar, Rugmini
; APPLICANT: Kim, Sang H
; TITLE OF INVENTION: Method for Producing Amide Compounds Using a Nitrile
; TITLE OF INVENTION: Hydratase from a Thermophilic Bacillus
; FILE REFERENCE: MSU 4.1-401
; CURRENT APPLICATION NUMBER: US/09/248,528C
; CURRENT FILING DATE: 1999-02-10
; EARLIER APPLICATION NUMBER: 60/083,485
; EARLIER FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO: 11
; LENGTH: 229
; TYPE: PRT
; ORGANISM: BR449
US-09-248-528-11

Query Match 83.9%; Score 26; DB 2; Length 229;
Best Local Similarity 83.3%; Pred. No. 4.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FSRYA 6
Db 161 FPRYAR 166

RESULT 34

US-09-549-108-11
; Sequence 11, Application US/09549108
; Patent No. 6214603
; GENERAL INFORMATION:
; APPLICANT: Oriol, Patrick J

; APPLICANT: Padmakumar, Rugmini
; APPLICANT: Kim, Sang H
; TITLE OF INVENTION: Method for Producing Amide Compounds Using a Nitrile
; TITLE OF INVENTION: Hydratase from a Thermophilic Bacillus
; FILE REFERENCE: MSU 4.1-486
; CURRENT APPLICATION NUMBER: US/09/549,108
; CURRENT FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/083,485
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 09/248,528
; PRIOR FILING DATE: 1999-02-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO: 11
; LENGTH: 229
; TYPE: PRT
; ORGANISM: BR449
US-09-549-108-11

Query Match 83.9%; Score 26; DB 2; Length 229;
Best Local Similarity 83.3%; Pred. No. 4.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FSRYA 6
Db 161 FPRYAR 166

RESULT 35

US-09-549-111-11
; Sequence 11, Application US/09549111
; Patent No. 6228633
; GENERAL INFORMATION:
; APPLICANT: Oriol, Patrick J
; APPLICANT: Padmakumar, Rugmini
; APPLICANT: Kim, Sang H
; TITLE OF INVENTION: Method for Producing Amide Compounds Using a Nitrile
; TITLE OF INVENTION: Hydratase from a Thermophilic Bacillus
; FILE REFERENCE: MSU 4.1-489
; CURRENT APPLICATION NUMBER: US/09/549,111
; CURRENT FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/083,485
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 09/248,528
; PRIOR FILING DATE: 1999-02-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO: 11
; LENGTH: 229
; TYPE: PRT
; ORGANISM: BR449
US-09-549-111-11

Query Match 83.9%; Score 26; DB 2; Length 229;
Best Local Similarity 83.3%; Pred. No. 4.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FSRYA 6
Db 161 FPRYAR 166

RESULT 36

US-09-549-106-11
; Sequence 11, Application US/09549106
; Patent No. 6242242
; GENERAL INFORMATION:
; APPLICANT: Oriol, Patrick J
; APPLICANT: Padmakumar, Rugmini
; APPLICANT: Kim, Sang H
; TITLE OF INVENTION: Method for Producing Amide Compounds Using a Nitrile
; TITLE OF INVENTION: Hydratase from a Thermophilic Bacillus
; FILE REFERENCE: MSU 4.1-487

```
; CURRENT APPLICATION NUMBER: US/09/549,106
; CURRENT FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/083,485
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 09/248,528
; PRIOR FILING DATE: 1999-02-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 11
; LENGTH: 229
; TYPE: PRT
; ORGANISM: BR449
US-09-549-106-11

Query Match      83.9%; Score 26; DB 2; Length 229;
Best Local Similarity 83.3%; Pred. No. 4.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 FSRVAR 6
        |||||
Db      161 FPRVAR 166

RESULT 37
US-09-550-394-11
; Sequence 11, Application US/09550394
; Patent No. 6287828
; GENERAL INFORMATION:
; APPLICANT: Oriol, Patrick J
; APPLICANT: Padmakumar, Rugmini
; APPLICANT: Kim, Sang H
; TITLE OF INVENTION: Method for Producing Amide Compounds Using a Nitrite
; TITLE OF INVENTION: Hydrtase from a Thermophilic Bacillus
; FILE REFERENCE: MSU 4.1-488
; CURRENT APPLICATION NUMBER: US/09/550,394
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/083,485
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 09/248,528
; PRIOR FILING DATE: 1999-02-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 11
; LENGTH: 229
; TYPE: PRT
; ORGANISM: BR449
US-09-550-394-11

Query Match      83.9%; Score 26; DB 2; Length 229;
Best Local Similarity 83.3%; Pred. No. 4.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 FSRVAR 6
        |||||
Db      161 FPRVAR 166

RESULT 38
US-09-489-039A-12718
; Sequence 12718, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Bretton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709-2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12718
; LENGTH: 233
```

```
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12718

Query Match      83.9%; Score 26; DB 2; Length 233;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 FSRYA 5
        |||||
Db      138 FSRYA 142

RESULT 39
US-09-949-016-6472
; Sequence 6472, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6472
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6472

Query Match      83.9%; Score 26; DB 2; Length 252;
Best Local Similarity 83.3%; Pred. No. 5.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 FSRVAR 6
        |||||
Db      240 FSRVAR 245

RESULT 40
US-09-949-016-9375
; Sequence 9375, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9375
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9375

Query Match      83.9%; Score 26; DB 2; Length 253;
```

Best Local Similarity 83.3%; Pred. No. 5.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FSRYAR 6
Db 241 FSRYAR 246

RESULT 41
US-09-198-452A-266
; Sequence 266, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Grifais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 266
; LENGTH: 264
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-266

Query Match 83.9%; Score 26; DB 2; Length 264;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRYA 5
Db 150 FSRYA 154

RESULT 42
US-09-438-185A-256
; Sequence 256, Application US/09438185A
; Patent No. 6822071
; GENERAL INFORMATION:
; APPLICANT: Stephens, Richard
; APPLICANT: Mitchell, Wayne
; APPLICANT: Kaiman, Sue
; APPLICANT: Davis, Ronald
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence
; FILE REFERENCE: 018941-000411US
; CURRENT APPLICATION NUMBER: US/09/438,185A
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/108,279
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 60/128,606
; PRIOR FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 1074
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 256
; LENGTH: 270
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; OTHER INFORMATION: CPH0254
US-09-438-185A-256

Query Match 83.9%; Score 26; DB 2; Length 270;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRYA 5
Db 150 FSRYA 154

RESULT 43
US-09-252-991A-25159
; Sequence 25159, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25159
; LENGTH: 276
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25159

Query Match 83.9%; Score 26; DB 2; Length 276;
Best Local Similarity 83.3%; Pred. No. 5.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FSRYAR 6
Db 71 FHRYAR 76

RESULT 44
US-09-252-991A-20485
; Sequence 20485, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20485
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20485

Query Match 83.9%; Score 26; DB 2; Length 283;
Best Local Similarity 83.3%; Pred. No. 5.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRYAR 6
Db 64 WSRYAR 69

RESULT 45
US-08-944-483-66
; Sequence 66, Application US/08944483
; Patent No. 6232456
; GENERAL INFORMATION:
; APPLICANT: COHEN, MAURICE
; APPLICANT: COLPITTS, TRACEY L.
; APPLICANT: FRIEDMAN, PAULA N.
; APPLICANT: GRANDOS, EDWARD N.
; APPLICANT: KLAS, MICHAEL R.
; APPLICANT: RUSSELL, JOHN C.

APPLICANT: STEWART, KENT D.
APPLICANT: STROPE, STEVEN D.
TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS
TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES
TITLE OF INVENTION: OF THE PROSTATE
NUMBER OF SEQUENCES: 76
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/944,483
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6183.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 66:
SEQUENCE CHARACTERISTICS:
LENGTH: 299 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6232456e
US-08-944-483-66

Query Match 83.9%; Score 26; DB 2; Length 299;
Best Local Similarity 83.3%; Pred. No. 6.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FSRYAR 6
Db 102 FSRYIR 107

RESULT 46
US-09-248-796A-14594
Sequence 14594, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Keith Weinstein et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
CURRENT FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 14594
LENGTH: 308
TYPE: PRT
ORGANISM: Candida albicans
US-09-248-796A-14594

Query Match 83.9%; Score 26; DB 2; Length 308;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSRYA 5
Db 190 FSRYIA 194

RESULT 47
US-09-386-642-12
Sequence 12, Application US/09386642
Patent No. 6420157
GENERAL INFORMATION:
APPLICANT: Darrow, Andrew
APPLICANT: Qi, Jenson
APPLICANT: Andrade-Gordon, Patricia
TITLE OF INVENTION: Zymogen Activation System
FILE REFERENCE: ORT-1028
CURRENT APPLICATION NUMBER: US/09/386,642
CURRENT FILING DATE: 1999-08-31
NUMBER OF SEQ ID NOS: 60
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 12
LENGTH: 319
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Fusion gene
US-09-386-642-12

Query Match 83.9%; Score 26; DB 2; Length 319;
Best Local Similarity 83.3%; Pred. No. 6.5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FSRYAR 6
Db 144 FSRYIR 149

RESULT 48
US-09-489-039A-14005
Sequence 14005, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 14005
LENGTH: 327
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-14005

Query Match 83.9%; Score 26; DB 2; Length 327;
Best Local Similarity 83.3%; Pred. No. 6.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSRYAR 6
Db 251 FSRYAR 256

RESULT 49
US-09-386-642-11
Sequence 11, Application US/09386642

```
; Patent No. 6420157
; GENERAL INFORMATION:
; APPLICANT: Darrow, Andrew
; APPLICANT: Qi, Jenson
; APPLICANT: Andrade-Gordon, Patricia
; TITLE OF INVENTION: Zymogen Activation System
; FILE REFERENCE: ORT-1028
; CURRENT APPLICATION NUMBER: US/09/386,642
; CURRENT FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 328
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Fusion gene
; OTHER INFORMATION: with homo sapien serine protease catalytic domain
US-09-386-642-11
```

```
Query Match      83.9%; Score 26; DB 2; Length 328;
Best Local Similarity 83.3%; Pred. No. 6.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY      1 FSRYAR 6
        |||||
Db      153 FSRYLR 158
```

```
RESULT 50
US-09-438-185A-960
; Sequence 960, Application US/09438185A
; Patent No. 6822071
; GENERAL INFORMATION:
; APPLICANT: Stephens, Richard
; APPLICANT: Mitchell, Wayne
; APPLICANT: Kalman, Sue
; APPLICANT: Davis, Ronald
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence
; FILE REFERENCE: 018941-000411US
; CURRENT APPLICATION NUMBER: US/09/438,185A
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/108,279
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 60/128,606
; PRIOR FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 1074
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 960
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; OTHER INFORMATION: CPN0958
US-09-438-185A-960
```

```
Query Match      83.9%; Score 26; DB 2; Length 337;
Best Local Similarity 83.3%; Pred. No. 6.8e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY      1 FSRYAR 6
        |||||
Db      9 FSRYLR 14
```

```
Search completed: August 29, 2006, 06:17:52
Job time : 14.3366 secs
```

THIS PAGE LEFT BLANK

GenCore version 5.1.9
Copyright (c) 1993 - 2006 BioCeleration Ltd.

OM protein - protein search, using sw model1

Run on: August 29, 2006, 09:58:08 ; Search time 72.5941 Seconds
(Without alignments)
38.285 Million cell updates/sec

Title: US-10-541-343-2

Perfect score: 31

Sequence: 1 PSRYAR 6

Scoring table:

BLOSUM62
Gap 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database : Published Applications AA Main:*

1: /EMC_Celerra_SIDS3/ptodata/2/pubppaa/US07_PUBCOMB pep: *
2: /EMC_Celerra_SIDS3/ptodata/2/pubppaa/US08_PUBCOMB pep: *
3: /EMC_Celerra_SIDS3/ptodata/2/pubppaa/US09_PUBCOMB pep: *
4: /EMC_Celerra_SIDS3/ptodata/2/pubppaa/US10A_PUBCOMB pep: *
5: /EMC_Celerra_SIDS3/ptodata/2/pubppaa/US10B_PUBCOMB pep: *
6: /EMC_Celerra_SIDS3/ptodata/2/pubppaa/US11_PUBCOMB pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	100.0	6	US-10-541-343-2	Sequence 2, Appli
2	31	100.0	10	US-10-541-343-4	Sequence 4, Appli
3	31	100.0	11	US-10-541-343-5	Sequence 5, Appli
4	31	100.0	13	US-10-541-343-6	Sequence 6, Appli
5	31	100.0	21	US-10-541-343-7	Sequence 7, Appli
6	31	100.0	33	US-09-989-481-4	Sequence 4, Appli
7	31	100.0	33	US-10-858-545-4	Sequence 4, Appli
8	31	100.0	33	US-10-858-266-4	Sequence 4, Appli
9	31	100.0	33	US-10-858-543-4	Sequence 4, Appli
10	31	100.0	33	US-10-858-144-4	Sequence 4, Appli
11	31	100.0	33	US-10-541-343-1	Sequence 1, Appli
12	31	100.0	125	US-11-188-298-11746	Sequence 11746, A
13	31	100.0	357	US-10-312-273-107	Sequence 977, App
14	31	100.0	357	US-10-289-762-977	Sequence 977, App
15	31	100.0	357	US-10-282-122A-54956	Sequence 54956, A
16	31	100.0	412	US-10-437-963-176098	Sequence 176098, A
17	31	100.0	463	US-10-369-493-23258	Sequence 23258, A
18	31	100.0	463	US-11-087-099-1159	Sequence 1159, Ap
19	31	100.0	463	US-11-188-298-12707	Sequence 12707, A
20	31	100.0	474	US-09-815-242-10270	Sequence 10270, A
21	31	100.0	474	US-10-369-493-23518	Sequence 23518, A
22	31	100.0	1225	US-10-450-763-54803	Sequence 54803, A
23	31	100.0	1258	US-10-450-763-55178	Sequence 55178, A
24	28	90.3	67	US-10-424-599-186222	Sequence 186222, A
25	28	90.3	79	US-10-424-599-146582	Sequence 146582, A
26	28	90.3	85	US-10-424-599-265564	Sequence 265564, A
27	28	90.3	90	US-10-183-687-178	Sequence 178, App

28	28	90.3	91	US-10-425-115-219020	Sequence 219020, A
29	28	90.3	123	US-10-767-701-51064	Sequence 51064, A
30	28	90.3	124	US-10-767-701-43382	Sequence 43382, A
31	28	90.3	132	US-10-183-687-174	Sequence 174, App
32	28	90.3	145	US-10-501-282-478	Sequence 478, App
33	28	90.3	172	US-10-501-282-480	Sequence 480, App
34	28	90.3	174	US-09-984-245-156	Sequence 156, App
35	28	90.3	174	US-09-966-262-156	Sequence 156, App
36	28	90.3	174	US-09-983-966-156	Sequence 156, App
37	28	90.3	174	US-10-059-395-156	Sequence 156, App
38	28	90.3	174	US-10-143-090-156	Sequence 156, App
39	28	90.3	174	US-10-960-251-156	Sequence 156, App
40	28	90.3	179	US-10-437-963-197182	Sequence 197182, A
41	28	90.3	225	US-11-096-568A-9600	Sequence 9600, Ap
42	28	90.3	234	US-10-425-115-314441	Sequence 314441, A
43	28	90.3	239	US-10-424-599-254610	Sequence 254644, A
44	28	90.3	244	US-10-987-454-122	Sequence 122, App
45	28	90.3	244	US-10-183-687-160	Sequence 160, App
46	28	90.3	256	US-10-425-114-54567	Sequence 54567, A
47	28	90.3	285	US-10-437-963-133297	Sequence 133297, A
48	28	90.3	301	US-10-987-454-127	Sequence 127, App
49	28	90.3	301	US-10-987-454-128	Sequence 128, App
50	28	90.3	301	US-10-450-763-37522	Sequence 37522, A
51	28	90.3	301	US-11-096-568A-9599	Sequence 9599, Ap
52	28	90.3	313	US-10-425-115-314430	Sequence 314430, A
53	28	90.3	348	US-10-424-599-240408	Sequence 240408, A
54	28	90.3	348	US-11-096-568A-9598	Sequence 9598, Ap
55	28	90.3	475	US-11-087-099-1900	Sequence 1900, Ap
56	28	90.3	479	US-11-087-099-8697	Sequence 8697, Ap
57	28	90.3	499	US-10-450-763-58083	Sequence 58083, A
58	28	90.3	520	US-10-450-763-60040	Sequence 60040, A
59	28	90.3	531	US-10-437-963-197185	Sequence 197185, A
60	28	90.3	538	US-10-437-963-107875	Sequence 107875, A
61	28	90.3	681	US-10-369-493-838	Sequence 838, App
62	28	90.3	749	US-10-437-963-133292	Sequence 133292, A
63	28	90.3	826	US-10-450-763-34395	Sequence 34395, A
64	28	90.3	1206	US-10-467-657-72	Sequence 72, Appl
65	28	90.3	1206	US-10-467-657-3892	Sequence 3892, Ap
66	28	90.3	2151	US-10-467-481-4	Sequence 4, Appli
67	27	87.1	43	US-10-437-963-140664	Sequence 140664, A
68	27	87.1	48	US-10-424-599-271620	Sequence 271620, A
69	27	87.1	51	US-10-724-972A-6411	Sequence 6411, Ap
70	27	87.1	63	US-10-424-599-258659	Sequence 258659, A
71	27	87.1	69	US-10-425-115-296787	Sequence 296787, A
72	27	87.1	74	US-10-424-599-258661	Sequence 258661, A
73	27	87.1	107	US-10-767-701-38384	Sequence 38384, A
74	27	87.1	190	US-10-424-599-202305	Sequence 202305, A
75	27	87.1	199	US-10-424-599-259517	Sequence 259517, A
76	27	87.1	202	US-10-424-599-187106	Sequence 187106, A
77	27	87.1	209	US-10-733-930-6140	Sequence 6140, Ap
78	27	87.1	220	US-09-925-301-1209	Sequence 1209, Ap
79	27	87.1	220	US-10-733-923-13264	Sequence 13264, A
80	27	87.1	220	US-10-733-923-13265	Sequence 13265, A
81	27	87.1	222	US-10-425-114-53064	Sequence 53064, A
82	27	87.1	272	US-10-437-963-107057	Sequence 107057, A
83	27	87.1	276	US-10-424-599-258658	Sequence 258658, A
84	27	87.1	289	US-10-424-599-16564	Sequence 16564, A
85	27	87.1	289	US-11-087-099-16564	Sequence 16564, A
86	27	87.1	292	US-11-188-298-14447	Sequence 14447, A
87	27	87.1	322	US-11-188-298-112237	Sequence 112237, A
88	27	87.1	322	US-10-437-963-177129	Sequence 177122, A
89	27	87.1	331	US-10-424-599-177129	Sequence 177122, A
90	27	87.1	338	US-10-437-963-103979	Sequence 103979, A
91	27	87.1	341	US-11-188-298-12428	Sequence 12428, A
92	27	87.1	354	US-09-815-242-11794	Sequence 11794, A
93	27	87.1	354	US-10-282-122A-66194	Sequence 66194, A
94	27	87.1	354	US-10-282-122A-69242	Sequence 69242, A
95	27	87.1	354	US-11-188-298-2415	Sequence 2415, Ap
96	27	87.1	354	US-11-188-298-3950	Sequence 3950, Ap
97	27	87.1	354	US-11-188-298-7657	Sequence 7657, Ap
98	27	87.1	374	US-10-437-963-102652	Sequence 102652, A
99	27	87.1	409	US-10-369-493-20350	Sequence 20350, A
100	27	87.1	410	US-10-369-493-1284	Sequence 1284, Ap

101	27	87.1	410	4	US-10-369-493-21624	Sequence 21624, A	174	27	87.1	676	5	US-10-732-923-22281	Sequence 22281, A
102	27	87.1	412	5	US-10-732-923-22297	Sequence 22297, A	175	27	87.1	1037	4	US-10-437-963-132655	Sequence 132655,
103	27	87.1	414	5	US-10-511-698-37	Sequence 37, App1	176	27	87.1	837	4	US-10-437-963-137165	Sequence 137165,
104	27	87.1	420	4	US-10-425-115-365824	Sequence 365824, A	177	26	83.9	1027	10	US-10-031-874A-98	Sequence 98, App1
105	27	87.1	425	5	US-10-732-923-22291	Sequence 22291, A	178	26	83.9	10	4	US-10-031-874A-101	Sequence 101, App
106	27	87.1	428	5	US-10-511-698-38	Sequence 38, App1	179	26	83.9	39	4	US-10-424-599-153126	Sequence 153126,
107	27	87.1	445	4	US-10-097-111-289	Sequence 289, App	180	26	83.9	42	4	US-10-425-115-359594	Sequence 359594,
108	27	87.1	453	4	US-10-369-493-14065	Sequence 14065, A	181	26	83.9	42	4	US-10-467-657-2278	Sequence 2278, Ap
109	27	87.1	456	4	US-10-282-122A-74836	Sequence 74836, A	182	26	83.9	50	4	US-10-424-599-263997	Sequence 263997,
110	27	87.1	456	4	US-10-282-122A-75400	Sequence 75400, A	183	26	83.9	50	4	US-10-428-275-86	Sequence 86, App1
111	27	87.1	456	6	US-11-087-009-5433	Sequence 5433, Ap	184	26	83.9	66	4	US-10-424-599-184533	Sequence 184533,
112	27	87.1	456	6	US-11-087-009-6176	Sequence 6176, Ap	185	26	83.9	66	4	US-10-425-115-209804	Sequence 209804,
113	27	87.1	456	6	US-11-188-298-16004	Sequence 16004, A	186	26	83.9	72	4	US-10-424-599-265076	Sequence 265076,
114	27	87.1	457	4	US-10-369-493-729	Sequence 729, App	187	26	83.9	82	4	US-10-424-599-274595	Sequence 274595,
115	27	87.1	457	4	US-10-282-122A-43084	Sequence 43084, A	188	26	83.9	89	4	US-10-767-701-52537	Sequence 52537, A
116	27	87.1	457	4	US-10-282-122A-60097	Sequence 60097, A	189	26	83.9	104	3	US-09-764-848-58	Sequence 28, App1
117	27	87.1	457	6	US-11-087-009-2494	Sequence 2494, Ap	190	26	83.9	104	3	US-09-764-848-58	Sequence 1132, Ap
118	27	87.1	457	6	US-11-087-009-3478	Sequence 3478, Ap	191	26	83.9	104	4	US-10-116-016-28	Sequence 28, App1
119	27	87.1	457	6	US-11-188-298-2380	Sequence 2380, Ap	192	26	83.9	104	4	US-10-222-020-28	Sequence 28, App1
120	27	87.1	457	6	US-11-188-298-2380	Sequence 2380, Ap	193	26	83.9	108	4	US-10-424-599-173271	Sequence 173271,
121	27	87.1	458	6	US-11-087-009-9639	Sequence 9639, Ap	194	26	83.9	108	6	US-11-087-009-4881	Sequence 4881, A
122	27	87.1	458	6	US-11-087-009-11216	Sequence 11216, A	195	26	83.9	109	6	US-11-096-568A-13389	Sequence 13389, Ap
123	27	87.1	458	6	US-11-188-298-8947	Sequence 8947, Ap	196	26	83.9	113	4	US-10-437-963-128432	Sequence 128432,
124	27	87.1	458	6	US-11-188-298-10343	Sequence 10343, A	197	26	83.9	119	6	US-11-049-536-370	Sequence 370, App
125	27	87.1	460	4	US-10-289-762-638	Sequence 762, App	198	26	83.9	119	6	US-11-199-739-370	Sequence 370, App
126	27	87.1	463	3	US-09-815-242-13918	Sequence 13918, A	199	26	83.9	120	6	US-11-049-536-518	Sequence 518, App
127	27	87.1	463	4	US-10-369-493-314	Sequence 314, App	200	26	83.9	124	6	US-11-199-739-518	Sequence 518, App
128	27	87.1	463	4	US-10-282-122A-78313	Sequence 78313, A	201	26	83.9	121	4	US-10-415-478A-604	Sequence 64, App1
129	27	87.1	463	6	US-11-087-009-7201	Sequence 7201, Ap	202	26	83.9	121	5	US-10-472-928-094	Sequence 2094, Ap
130	27	87.1	463	6	US-11-087-009-11335	Sequence 11335, A	203	26	83.9	125	3	US-09-828-708-14	Sequence 14, App1
131	27	87.1	463	6	US-11-188-298-6612	Sequence 6612, Ap	204	26	83.9	125	4	US-10-425-115-283844	Sequence 283844,
132	27	87.1	463	6	US-11-188-298-21475	Sequence 21475, A	205	26	83.9	125	5	US-10-630-009-14	Sequence 14, App1
133	27	87.1	472	6	US-11-087-009-13391	Sequence 12391, A	206	26	83.9	127	4	US-10-714-353-6	Sequence 6, App11
134	27	87.1	472	6	US-11-188-298-11371	Sequence 11371, A	207	26	83.9	128	4	US-10-424-599-155994	Sequence 155994,
135	27	87.1	478	4	US-10-369-493-16830	Sequence 16830, A	208	26	83.9	130	6	US-10-118-100-51	Sequence 51, App1
136	27	87.1	478	6	US-11-188-298-18308	Sequence 18308, A	209	26	83.9	130	6	US-10-109-264-51	Sequence 51, App1
137	27	87.1	483	4	US-10-369-493-14243	Sequence 14243, A	210	26	83.9	132	4	US-10-767-701-43324	Sequence 43324, A
138	27	87.1	484	6	US-10-026-140-2	Sequence 2, App11	211	26	83.9	135	4	US-10-425-115-10389	Sequence 210389,
139	27	87.1	484	6	US-11-147-725-2	Sequence 2, App11	212	26	83.9	137	4	US-10-479-670-108	Sequence 108, App
140	27	87.1	485	6	US-11-087-009-9429	Sequence 9429, Ap	213	26	83.9	137	4	US-10-425-115-335852	Sequence 335852,
141	27	87.1	485	6	US-11-188-298-19761	Sequence 19761, A	214	26	83.9	140	4	US-10-006-773-4	Sequence 4, App11
142	27	87.1	488	3	US-09-981-900B-7	Sequence 7, App11	215	26	83.9	145	5	US-10-467-657-4245	Sequence 42456, Ap
143	27	87.1	491	4	US-10-282-122A-69578	Sequence 69578, A	216	26	83.9	151	4	US-10-425-115-22327	Sequence 22327, App
144	27	87.1	502	4	US-10-156-761-11329	Sequence 11329, A	217	26	83.9	169	4	US-10-106-698-6779	Sequence 6779, Ap
145	27	87.1	505	4	US-10-369-493-14505	Sequence 14505, A	218	26	83.9	174	4	US-10-369-493-2230	Sequence 2230, Ap
146	27	87.1	505	4	US-10-369-493-14906	Sequence 14906, A	219	26	83.9	185	4	US-10-425-115-113735	Sequence 317325,
147	27	87.1	538	4	US-10-724-972A-4824	Sequence 4824, Ap	220	26	83.9	189	5	US-10-467-657-4524	Sequence 4524, Ap
148	27	87.1	541	4	US-10-424-599-171116	Sequence 171116, A	221	26	83.9	198	5	US-10-450-753-32783	Sequence 32783, A
149	27	87.1	545	4	US-10-425-114-45091	Sequence 45091, A	222	26	83.9	211	4	US-10-424-599-240815	Sequence 240815,
150	27	87.1	579	3	US-09-824-588-2	Sequence 2, App11	223	26	83.9	212	4	US-10-437-963-152547	Sequence 152547,
151	27	87.1	579	4	US-10-359-289-2	Sequence 2, App11	224	26	83.9	219	6	US-11-096-568A-16580	Sequence 16580, A
152	27	87.1	592	5	US-10-408-713-8	Sequence 8, App11	225	26	83.9	219	6	US-11-096-568A-25155	Sequence 25155, A
153	27	87.1	592	5	US-10-732-923-22115	Sequence 22115, A	226	26	83.9	226	4	US-10-425-115-220678	Sequence 220678,
154	27	87.1	592	5	US-10-732-923-22125	Sequence 22125, A	227	26	83.9	231	3	US-09-898-837A-43	Sequence 43, App1
155	27	87.1	607	4	US-10-369-493-11846	Sequence 11846, A	228	26	83.9	231	3	US-10-424-599-158028	Sequence 158028,
156	27	87.1	607	4	US-10-369-493-12204	Sequence 12204, A	229	26	83.9	234	3	US-09-898-837A-48	Sequence 48, App1
157	27	87.1	608	4	US-10-369-493-12230	Sequence 12230, A	230	26	83.9	241	4	US-10-156-761-14640	Sequence 14640, A
158	27	87.1	607	4	US-10-369-493-11543	Sequence 11543, A	231	26	83.9	244	4	US-10-425-115-275008	Sequence 275008,
159	27	87.1	630	3	US-09-738-625-6059	Sequence 6059, Ap	232	26	83.9	247	3	US-09-880-748-878	Sequence 878, App
160	27	87.1	630	4	US-10-627-476-548	Sequence 548, App	233	26	83.9	247	3	US-09-880-748-1090	Sequence 1090, Ap
161	27	87.1	635	5	US-10-494-677-298	Sequence 298, App	234	26	83.9	247	3	US-09-880-748-1099	Sequence 1099, Ap
162	27	87.1	647	5	US-10-732-923-22114	Sequence 22114, A	235	26	83.9	247	3	US-09-880-748-1328	Sequence 1328, Ap
163	27	87.1	648	5	US-10-732-923-22116	Sequence 22116, A	236	26	83.9	247	3	US-10-293-418-878	Sequence 878, App
164	27	87.1	650	4	US-10-408-713-4	Sequence 4, App11	237	26	83.9	247	4	US-10-293-418-1099	Sequence 1099, Ap
165	27	87.1	650	4	US-10-408-713-10	Sequence 10, App1	238	26	83.9	247	4	US-10-293-418-1099	Sequence 1099, Ap
166	27	87.1	650	5	US-10-732-923-22282	Sequence 22282, A	239	26	83.9	247	4	US-10-293-418-1328	Sequence 1328, Ap
167	27	87.1	650	5	US-10-732-923-22283	Sequence 22283, A	240	26	83.9	247	6	US-11-054-515-878	Sequence 878, App
168	27	87.1	650	5	US-10-732-923-22284	Sequence 22284, A	241	26	83.9	247	6	US-11-054-515-1090	Sequence 1090, Ap
169	27	87.1	650	5	US-10-732-923-22285	Sequence 22285, A	242	26	83.9	247	6	US-11-054-515-1099	Sequence 1099, Ap
170	27	87.1	650	5	US-10-732-923-22294	Sequence 22294, A	243	26	83.9	247	6	US-11-054-515-1328	Sequence 1328, Ap
171	27	87.1	650	5	US-10-756-149-4983	Sequence 4983, Ap	244	26	83.9	247	6	US-11-266-444-878	Sequence 878, App
172	27	87.1	650	5	US-10-511-698-36	Sequence 36, App1	245	26	83.9	247	6	US-11-266-444-1090	Sequence 1090, Ap
173	27	87.1	672	4	US-10-346-863-10	Sequence 10, App1	246	26	83.9	247	6	US-11-266-444-1099	Sequence 1099, Ap

247	26	83.9	247	6	US-11-266-444-1328	Sequence 1328, Ap	320	26	83.9	343	5	US-10-784-004-661	Sequence 661, App
248	26	83.9	249	3	US-09-815-242-10880	Sequence 10880, A	321	26	83.9	343	5	US-10-784-004-663	Sequence 663, App
249	26	83.9	250	3	US-11-087-099-1459	Sequence 1459, Ap	322	26	83.9	343	5	US-10-784-004-1060	Sequence 1060, App
250	26	83.9	251	3	US-09-925-301-1050	Sequence 1050, Ap	323	26	83.9	343	6	US-11-017-119-2	Sequence 2, Appl1
251	26	83.9	251	4	US-10-767-701-45303	Sequence 45303, A	324	26	83.9	343	6	US-11-050-926-262	Sequence 262, App
252	26	83.9	252	4	US-10-087-192-1242	Sequence 1242, Ap	325	26	83.9	343	6	US-11-080-591-86	Sequence 86, Appl1
253	26	83.9	253	3	US-09-880-748-1199	Sequence 1199, Ap	326	26	83.9	343	6	US-11-019-711-130	Sequence 130, App
254	26	83.9	254	4	US-10-293-418-1199	Sequence 1199, Ap	327	26	83.9	348	5	US-10-470-0488-72	Sequence 72, Appl
255	26	83.9	255	6	US-11-054-515-1199	Sequence 1199, Ap	328	26	83.9	349	5	US-10-739-930-9327	Sequence 9327, Ap
256	26	83.9	256	3	US-11-266-444-1199	Sequence 1199, Ap	329	26	83.9	350	4	US-10-424-599-188274	Sequence 182374, Ap
257	26	83.9	257	3	US-09-880-748-1012	Sequence 1012, Ap	330	26	83.9	356	4	US-10-369-493-3321	Sequence 3321, Ap
258	26	83.9	258	4	US-10-293-418-1012	Sequence 1012, Ap	331	26	83.9	355	3	US-09-995-493-22	Sequence 22, Appl1
259	26	83.9	259	6	US-11-054-515-1012	Sequence 1012, Ap	332	26	83.9	377	6	US-11-096-568A-16578	Sequence 16578, A
260	26	83.9	260	6	US-11-266-444-1012	Sequence 1012, Ap	333	26	83.9	386	4	US-10-042-865-32	Sequence 32, Appl
261	26	83.9	261	4	US-10-289-762-266	Sequence 266, App	334	26	83.9	393	5	US-10-821-234-1043	Sequence 1043, Ap
262	26	83.9	262	4	US-10-479-670-192	Sequence 192, App	335	26	83.9	395	4	US-10-437-963-15167	Sequence 15167, A
263	26	83.9	263	4	US-10-425-115-282322	Sequence 282322, A	336	26	83.9	407	5	US-10-732-923-17540	Sequence 17540, A
264	26	83.9	264	4	US-10-282-122A-57255	Sequence 57255, A	337	26	83.9	431	6	US-11-045-004-790	Sequence 790, Appl
265	26	83.9	265	4	US-10-437-963-132254	Sequence 132254, A	338	26	83.9	435	4	US-10-000-512-18	Sequence 18, Appl
266	26	83.9	266	4	US-10-425-115-10385	Sequence 210385, A	339	26	83.9	435	4	US-10-074-566-18	Sequence 18, Appl
267	26	83.9	267	4	US-10-425-115-210386	Sequence 210386, A	340	26	83.9	435	5	US-10-453-195-18	Sequence 18, Appl
268	26	83.9	268	6	US-11-087-099-2401	Sequence 2401, Ap	341	26	83.9	448	4	US-10-282-122A-72853	Sequence 72853, A
269	26	83.9	269	6	US-11-087-099-4667	Sequence 4667, Ap	342	26	83.9	450	4	US-10-369-493-19569	Sequence 19569, A
270	26	83.9	270	6	US-11-087-099-5728	Sequence 5728, Ap	343	26	83.9	451	4	US-10-428-275-84	Sequence 84, Appl
271	26	83.9	271	6	US-11-087-099-5782	Sequence 5782, Ap	344	26	83.9	451	4	US-10-428-275-100	Sequence 100, Appl
272	26	83.9	272	6	US-11-087-099-7730	Sequence 7730, Ap	345	26	83.9	451	4	US-10-428-275-102	Sequence 102, App
273	26	83.9	273	6	US-11-087-099-11152	Sequence 11152, Ap	346	26	83.9	452	6	US-11-188-298-19762	Sequence 19762, A
274	26	83.9	274	6	US-11-096-568A-23464	Sequence 23464, A	347	26	83.9	467	4	US-10-282-122A-69918	Sequence 69918, A
275	26	83.9	275	6	US-11-096-568A-25153	Sequence 25153, A	348	26	83.9	473	4	US-10-437-963-151470	Sequence 151470, A
276	26	83.9	276	6	US-11-087-099-5910	Sequence 5910, Ap	349	26	83.9	474	5	US-10-732-923-19590	Sequence 19590, A
277	26	83.9	277	4	US-10-425-114-50779	Sequence 50779, A	350	26	83.9	495	5	US-10-756-149-5454	Sequence 5454, A
278	26	83.9	278	6	US-11-188-298-13239	Sequence 13239, A	351	26	83.9	495	5	US-10-450-763-40077	Sequence 40077, A
279	26	83.9	279	4	US-10-437-963-166051	Sequence 166051, A	352	26	83.9	497	3	US-09-815-242-13344	Sequence 13344, A
280	26	83.9	280	4	US-10-087-192-1239	Sequence 1239, Ap	353	26	83.9	497	3	US-10-282-122A-73512	Sequence 73512, A
281	26	83.9	281	4	US-10-051-874-89	Sequence 89, Appl	354	26	83.9	497	5	US-10-472-928-1718	Sequence 1718, Ap
282	26	83.9	282	6	US-11-098-686-11075	Sequence 11075, A	355	26	83.9	499	4	US-10-156-761-11332	Sequence 11332, A
283	26	83.9	283	6	US-09-789-210-66	Sequence 66, Appl	356	26	83.9	499	6	US-11-188-298-5047	Sequence 5047, Ap
284	26	83.9	284	5	US-10-741-849-7146	Sequence 7146, Ap	357	26	83.9	504	5	US-10-617-320-3732	Sequence 3732, Ap
285	26	83.9	285	3	US-11-096-568A-23462	Sequence 23462, Ap	358	26	83.9	509	5	US-10-470-0488-62	Sequence 62, Appl
286	26	83.9	286	6	US-10-774-355A-1505	Sequence 1505, Ap	359	26	83.9	520	4	US-10-128-714-8505	Sequence 8505, Ap
287	26	83.9	287	4	US-10-168-663-15	Sequence 15, Appl	360	26	83.9	528	4	US-10-128-714-8505	Sequence 8505, Ap
288	26	83.9	288	4	US-10-424-599-243121	Sequence 243121, A	361	26	83.9	543	4	US-10-282-122A-59680	Sequence 59680, A
289	26	83.9	289	5	US-10-015-989A-12	Sequence 12, Appl	362	26	83.9	547	3	US-09-895-860-2	Sequence 2, Appl1
290	26	83.9	290	5	US-10-320-797-4121	Sequence 4121, Ap	363	26	83.9	547	3	US-10-377-072-2	Sequence 2, Appl1
291	26	83.9	291	5	US-10-015-989A-11	Sequence 11, Appl	364	26	83.9	547	4	US-10-377-072-2	Sequence 2, Appl1
292	26	83.9	292	3	US-09-764-875-844	Sequence 844, App	365	26	83.9	549	5	US-10-462-706-197	Sequence 197, App
293	26	83.9	293	3	US-09-833-245-191	Sequence 191, App	366	26	83.9	557	5	US-10-450-763-37783	Sequence 37783, Ap
294	26	83.9	294	6	US-11-188-298-11778	Sequence 11778, A	367	26	83.9	558	4	US-10-369-493-3398	Sequence 3398, Ap
295	26	83.9	295	6	US-11-264-096-191	Sequence 191, App	368	26	83.9	571	3	US-09-931-836-23	Sequence 23, Appl
296	26	83.9	296	4	US-10-109-616-2	Sequence 2, Appl1	369	26	83.9	571	3	US-09-833-245-1090	Sequence 1090, Ap
297	26	83.9	297	4	US-10-042-865-153	Sequence 153, App	370	26	83.9	571	4	US-10-036-342-23	Sequence 23, Appl
298	26	83.9	298	4	US-10-724-972A-5121	Sequence 5121, Ap	371	26	83.9	571	4	US-10-036-342-23	Sequence 23, Appl
299	26	83.9	299	4	US-10-051-874-86	Sequence 86, Appl	372	26	83.9	571	4	US-10-028-072-542	Sequence 542, App
300	26	83.9	300	4	US-10-051-874-87	Sequence 87, Appl	373	26	83.9	571	4	US-10-035-855-23	Sequence 23, Appl
301	26	83.9	301	4	US-10-042-865-151	Sequence 151, App	374	26	83.9	571	4	US-10-140-808-542	Sequence 542, App
302	26	83.9	302	4	US-10-042-865-152	Sequence 152, App	375	26	83.9	571	4	US-10-121-049-542	Sequence 542, App
303	26	83.9	303	4	US-10-042-865-154	Sequence 154, App	376	26	83.9	571	4	US-10-123-904-542	Sequence 542, App
304	26	83.9	304	4	US-10-425-114-36811	Sequence 36811, A	377	26	83.9	571	4	US-10-140-470-542	Sequence 542, App
305	26	83.9	305	4	US-10-037-417-133	Sequence 133, App	378	26	83.9	571	4	US-10-175-746-542	Sequence 542, App
306	26	83.9	306	4	US-10-037-417-134	Sequence 134, App	379	26	83.9	571	4	US-10-176-918-542	Sequence 542, App
307	26	83.9	307	5	US-10-784-004-341	Sequence 341, App	380	26	83.9	571	4	US-10-176-921-542	Sequence 542, App
308	26	83.9	308	5	US-10-784-004-342	Sequence 342, App	381	26	83.9	571	4	US-10-227-884-210	Sequence 210, App
309	26	83.9	309	5	US-10-784-004-317	Sequence 317, App	382	26	83.9	571	4	US-10-026-214-23	Sequence 23, Appl
310	26	83.9	310	6	US-11-019-711-133	Sequence 133, App	383	26	83.9	571	4	US-10-137-865-542	Sequence 542, App
311	26	83.9	311	6	US-11-019-711-134	Sequence 134, App	384	26	83.9	571	4	US-10-140-474-542	Sequence 542, App
312	26	83.9	312	6	US-09-948-094-2	Sequence 2, Appl1	385	26	83.9	571	4	US-10-035-719-23	Sequence 23, Appl1
313	26	83.9	313	4	US-10-176-847-86	Sequence 86, Appl	386	26	83.9	571	4	US-10-142-114-542	Sequence 542, App
314	26	83.9	314	4	US-10-097-340-262	Sequence 262, App	387	26	83.9	571	4	US-10-143-431-542	Sequence 542, App
315	26	83.9	315	4	US-10-074-566-125	Sequence 125, App	388	26	83.9	571	4	US-10-230-163-210	Sequence 210, App
316	26	83.9	316	4	US-10-042-865-150	Sequence 150, App	389	26	83.9	571	4	US-10-036-160-23	Sequence 23, Appl
317	26	83.9	317	4	US-10-037-417-130	Sequence 130, App	390	26	83.9	571	4	US-10-230-338-210	Sequence 210, App
318	26	83.9	318	4	US-10-311-591A-3	Sequence 3, Appl1	391	26	83.9	571	4	US-10-142-419-542	Sequence 542, App
319	26	83.9	319	4	US-10-712-124-8	Sequence 8, Appl1	392	26	83.9	571	4	US-10-218-631-210	Sequence 210, App

393	26	83.9	571	4	US-10-035-958-23	Sequence 23, App1
394	26	83.9	571	4	US-10-036-150-23	Sequence 23, App1
395	26	83.9	571	4	US-10-123-262-542	Sequence 542, App
396	26	83.9	571	4	US-10-143-423-542	Sequence 542, App
397	26	83.9	571	4	US-10-230-414-210	Sequence 210, App
398	26	83.9	571	4	US-10-121-050-542	Sequence 542, App
399	26	83.9	571	4	US-10-141-755-542	Sequence 542, App
400	26	83.9	571	4	US-10-143-033-542	Sequence 542, App
401	26	83.9	571	4	US-10-232-224-210	Sequence 210, App
402	26	83.9	571	4	US-10-123-108-542	Sequence 542, App
403	26	83.9	571	4	US-10-123-223-542	Sequence 542, App
404	26	83.9	571	4	US-10-123-261-542	Sequence 542, App
405	26	83.9	571	4	US-10-140-921-542	Sequence 542, App
406	26	83.9	571	4	US-10-140-921-542	Sequence 542, App
407	26	83.9	571	4	US-10-216-159A-210	Sequence 210, App
408	26	83.9	571	4	US-10-121-045-542	Sequence 542, App
409	26	83.9	571	4	US-10-123-222-542	Sequence 542, App
410	26	83.9	571	4	US-10-123-903-542	Sequence 542, App
411	26	83.9	571	4	US-10-124-819-542	Sequence 542, App
412	26	83.9	571	4	US-10-124-822-542	Sequence 542, App
413	26	83.9	571	4	US-10-140-925-542	Sequence 542, App
414	26	83.9	571	4	US-10-160-498-542	Sequence 542, App
415	26	83.9	571	4	US-10-218-849-210	Sequence 210, App
416	26	83.9	571	4	US-10-227-873-210	Sequence 210, App
417	26	83.9	571	4	US-10-227-883-210	Sequence 210, App
418	26	83.9	571	4	US-10-124-824-542	Sequence 542, App
419	26	83.9	571	4	US-10-127-825A-542	Sequence 542, App
420	26	83.9	571	4	US-10-127-829A-542	Sequence 542, App
421	26	83.9	571	4	US-10-127-835A-542	Sequence 542, App
422	26	83.9	571	4	US-10-127-839A-542	Sequence 542, App
423	26	83.9	571	4	US-10-127-901A-542	Sequence 542, App
424	26	83.9	571	4	US-10-128-693A-542	Sequence 542, App
425	26	83.9	571	4	US-10-131-813A-542	Sequence 542, App
426	26	83.9	571	4	US-10-131-818A-542	Sequence 542, App
427	26	83.9	571	4	US-10-131-823A-542	Sequence 542, App
428	26	83.9	571	4	US-10-131-824A-542	Sequence 542, App
429	26	83.9	571	4	US-10-131-830A-542	Sequence 542, App
430	26	83.9	571	4	US-10-131-837A-542	Sequence 542, App
431	26	83.9	571	4	US-10-137-872A-542	Sequence 542, App
432	26	83.9	571	4	US-10-147-500-542	Sequence 542, App
433	26	83.9	571	4	US-10-147-502-542	Sequence 542, App
434	26	83.9	571	4	US-10-147-515-542	Sequence 542, App
435	26	83.9	571	4	US-10-147-517-542	Sequence 542, App
436	26	83.9	571	4	US-10-147-526-542	Sequence 542, App
437	26	83.9	571	4	US-10-147-527-542	Sequence 542, App
438	26	83.9	571	4	US-10-121-041-542	Sequence 542, App
439	26	83.9	571	4	US-10-121-043-542	Sequence 542, App
440	26	83.9	571	4	US-10-121-047-542	Sequence 542, App
441	26	83.9	571	4	US-10-123-215-542	Sequence 542, App
442	26	83.9	571	4	US-10-123-902-542	Sequence 542, App
443	26	83.9	571	4	US-10-123-908-542	Sequence 542, App
444	26	83.9	571	4	US-10-123-909-542	Sequence 542, App
445	26	83.9	571	4	US-10-123-910-542	Sequence 542, App
446	26	83.9	571	4	US-10-124-813-542	Sequence 542, App
447	26	83.9	571	4	US-10-124-817-542	Sequence 542, App
448	26	83.9	571	4	US-10-125-922-542	Sequence 542, App
449	26	83.9	571	4	US-10-125-924-542	Sequence 542, App
450	26	83.9	571	4	US-10-140-860-542	Sequence 542, App
451	26	83.9	571	4	US-10-142-417-542	Sequence 542, App
452	26	83.9	571	4	US-10-147-519-542	Sequence 542, App
453	26	83.9	571	4	US-10-157-782-542	Sequence 542, App
454	26	83.9	571	4	US-10-152-395-542	Sequence 542, App
455	26	83.9	571	4	US-10-219-076-210	Sequence 210, App
456	26	83.9	571	4	US-10-230-434-210	Sequence 210, App
457	26	83.9	571	4	US-10-125-926A-542	Sequence 542, App
458	26	83.9	571	4	US-10-125-930A-542	Sequence 542, App
459	26	83.9	571	4	US-10-127-831A-542	Sequence 542, App
460	26	83.9	571	4	US-10-127-837A-542	Sequence 542, App
461	26	83.9	571	4	US-10-127-838A-542	Sequence 542, App
462	26	83.9	571	4	US-10-127-842A-542	Sequence 542, App
463	26	83.9	571	4	US-10-127-843A-542	Sequence 542, App
464	26	83.9	571	4	US-10-127-845A-542	Sequence 542, App
465	26	83.9	571	4	US-10-127-846A-542	Sequence 542, App

466

26

83.9

571

4

US-10-127-848A-542

Sequence 542, App

467

26

83.9

571

4

US-10-127-849A-542

Sequence 542, App

468

26

83.9

571

4

US-10-127-850A-542

Sequence 542, App

469

26

83.9

571

4

US-10-127-851A-542

Sequence 542, App

470

26

83.9

571

4

US-10-128-664A-542

Sequence 542, App

471

26

83.9

571

4

US-10-128-666A-542

Sequence 542, App

472

26

83.9

571

4

US-10-128-690A-542

Sequence 542, App

473

26

83.9

571

4

US-10-128-691A-542

Sequence 542, App

474

26

83.9

571

4

US-10-131-819A-542

Sequence 542, App

475

26

83.9

571

4

US-10-131-829A-542

Sequence 542, App

476

26

83.9

571

4

US-10-131-836A-542

Sequence 542, App

477

26

83.9

571

4

US-10-146-729-542

Sequence 542, App

478

26

83.9

571

4

US-10-146-791-542

Sequence 542, App

479

26

83.9

571

4

US-10-147-484-542

Sequence 542, App

480

26

83.9

571

4

US-10-147-508-542

Sequence 542, App

481

26

83.9

571

4

US-10-147-512-542

Sequence 542, App

482

26

83.9

571

4

US-10-175-735-542

Sequence 542, App

483

26

83.9

571

4

US-10-121-040-542

Sequence 542, App

484

26

83.9

571

4

US-10-121-056-542

Sequence 542, App

485

26

83.9

571

4

US-10-121-061-542

Sequence 542, App

486

26

83.9

571

4

US-10-123-235-542

Sequence 542, App

487

26

83.9

571

4

US-10-124-818-542

Sequence 542, App

488

26

83.9

571

4

US-10-137-868-542

Sequence 542, App

489

26

83.9

571

4

US-10-147-492-542

Sequence 542, App

490

26

83.9

571

4

US-10-158-782-542

Sequence 542, App

491

26

83.9

571

4

US-10-123-907-542

Sequence 542, App

492

26

83.9

571

4

US-10-123-907-542

Sequence 542, App

493

26

83.9

571

4

US-10-124-815-542

Sequence 542, App

494

26

83.9

571

4

US-10-125-921A-542

Sequence 542, App

495

26

83.9

571

4

US-10-125-928A-542

Sequence 542, App

496

26

83.9

571

4

US-10-127-821A-542

Sequence 542, App

497

26

83.9

571

4

US-10-127-822A-542

Sequence 542, App

498

26

83.9

571

4

US-10-127-824A-542

Sequence 542, App

499

26

83.9

571

4

US-10-127-826A-542

Sequence 542, App

500

26

83.9

571

4

US-10-127-827A-542

Sequence 542, App

ALIGNMENTS

RESULT 1

US-10-541-343-2

Sequence 2, Application US/10541343

Publication No. US20060052299A1

GENERAL INFORMATION:

APPLICANT: Chau, Raymond Ming Wah

TITLE OF INVENTION: NMPF Peptides and Compositions and Methods of Use

FILE REFERENCE: 2001-10305

CURRENT APPLICATION NUMBER: US/10/541,343

CURRENT FILING DATE: 2005-07-05

PRIOR APPLICATION NUMBER: PCT/US2004/001468

PRIOR FILING DATE: 2004-01-21

PRIOR APPLICATION NUMBER: 60/441,722

PRIOR FILING DATE: 2003-01-21

NUMBER OF SEQ ID NOS: 7

SOFTWARE: PatentIn version 3.3

SEQ ID NO 2

LENGTH: 6

TYPE: PRT

ORGANISM: Artificial

FEATURE:

OTHER INFORMATION: Synthetic peptide

US-10-541-343-2

Query Match

Best Local Similarity

Matches

6;

Conservative

0;

Mismatches

0;

Gaps

0;

Indels

0;

Score 31;

DB 5;

Length 6;

Pred. No. 1.9e+06;

1 FSRVAR 6

1 FSRVAR 6

Db

```
RESULT 2
US-10-541-343-4
; Sequence 4, Application US/10541343
; Publication No. US20060052299A1
; GENERAL INFORMATION:
; APPLICANT: Chau, Raymond Ming Wah
; APPLICANT: Ko, Pui-Yuk Dorothy
; TITLE OF INVENTION: MNTF Peptides and Compositions and Methods of Use
; FILE REFERENCE: 2001-103US
; CURRENT APPLICATION NUMBER: US/10/541,343
; PRIOR FILING DATE: 2005-07-05
; PRIOR APPLICATION NUMBER: PCT/US2004/001468
; PRIOR FILING DATE: 2004-01-21
; PRIOR APPLICATION NUMBER: 60/441,722
; PRIOR FILING DATE: 2003-01-21
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
US-10-541-343-4

Query Match      100.0%; Score 31; DB 5; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.8;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 FSRYAR 6
      |||||
Db      5 FSRYAR 10

RESULT 3
US-10-541-343-5
; Sequence 5, Application US/10541343
; Publication No. US20060052299A1
; GENERAL INFORMATION:
; APPLICANT: Chau, Raymond Ming Wah
; APPLICANT: Ko, Pui-Yuk Dorothy
; TITLE OF INVENTION: MNTF Peptides and Compositions and Methods of Use
; FILE REFERENCE: 2001-103US
; CURRENT APPLICATION NUMBER: US/10/541,343
; PRIOR FILING DATE: 2005-07-05
; PRIOR APPLICATION NUMBER: PCT/US2004/001468
; PRIOR FILING DATE: 2004-01-21
; PRIOR APPLICATION NUMBER: 60/441,722
; PRIOR FILING DATE: 2003-01-21
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
US-10-541-343-5

Query Match      100.0%; Score 31; DB 5; Length 11;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 FSRYAR 6
      |||||
Db      1 FSRYAR 6

RESULT 4
US-10-541-343-6
; Sequence 6, Application US/10541343
; Publication No. US20060052299A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Chau, Raymond Ming Wah
; APPLICANT: Ko, Pui-Yuk Dorothy
; TITLE OF INVENTION: MNTF Peptides and Compositions and Methods of Use
; FILE REFERENCE: 2001-103US
; CURRENT APPLICATION NUMBER: US/10/541,343
; PRIOR FILING DATE: 2005-07-05
; PRIOR APPLICATION NUMBER: PCT/US2004/001468
; PRIOR FILING DATE: 2004-01-21
; PRIOR APPLICATION NUMBER: 60/441,722
; PRIOR FILING DATE: 2003-01-21
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 6
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
US-10-541-343-6

Query Match      100.0%; Score 31; DB 5; Length 13;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 FSRYAR 6
      |||||
Db      7 FSRYAR 12

RESULT 5
US-10-541-343-7
; Sequence 7, Application US/10541343
; Publication No. US20060052299A1
; GENERAL INFORMATION:
; APPLICANT: Chau, Raymond Ming Wah
; APPLICANT: Ko, Pui-Yuk Dorothy
; TITLE OF INVENTION: MNTF Peptides and Compositions and Methods of Use
; FILE REFERENCE: 2001-103US
; CURRENT APPLICATION NUMBER: US/10/541,343
; PRIOR FILING DATE: 2005-07-05
; PRIOR APPLICATION NUMBER: PCT/US2004/001468
; PRIOR FILING DATE: 2004-01-21
; PRIOR APPLICATION NUMBER: 60/441,722
; PRIOR FILING DATE: 2003-01-21
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 7
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
US-10-541-343-7

Query Match      100.0%; Score 31; DB 5; Length 21;
Best Local Similarity 100.0%; Pred. No. 7.9;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 FSRYAR 6
      |||||
Db      5 FSRYAR 10

RESULT 6
US-09-989-481-4
; Sequence 4, Application US/09989481
; Publication No. US20020086831A1
; GENERAL INFORMATION:
; APPLICANT: Chau, Raymond M.W.
; TITLE OF INVENTION: Isolation and Use of Motoneuronotropic Factors
; FILE REFERENCE: 12592-4
; CURRENT APPLICATION NUMBER: US/09/989,481
```

```
/ CURRENT FILING DATE: 2001-11-20
/ PRIOR APPLICATION NUMBER: US 09/633,447
/ PRIOR FILING DATE: 2000-08-07
/ PRIOR APPLICATION NUMBER: US 08/9928862
/ PRIOR FILING DATE: 1997-09-12
/ PRIOR APPLICATION NUMBER: US 08/751225
/ PRIOR FILING DATE: 1996-11-15
/ PRIOR APPLICATION NUMBER: US 60/026792
/ PRIOR FILING DATE: 1996-09-27
/ NUMBER OF SEQ ID NOS: 5
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 4
/ LENGTH: 33
/ TYPE: PRT
/ ORGANISM: Homo Sapiens
US-09-989-481-4

Query Match          100.0%; Score 31; DB 3; Length 33;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRVAR 6
   |||||
Db 17 FSRVAR 22

RESULT 7
US-10-858-545-4
/ Sequence 4, Application US/10858545
/ Publication No. US20040224894A1
/ GENERAL INFORMATION:
/ APPLICANT: Chau, Raymond M.W.
/ TITLE OF INVENTION: Isolation and Use of Motoneurontropic Factors
/ FILE REFERENCE: 12592-3
/ CURRENT APPLICATION NUMBER: US/10/858,545
/ CURRENT FILING DATE: 2004-06-01
/ PRIOR APPLICATION NUMBER: US/09/592,018
/ PRIOR FILING DATE: 2000-06-12
/ PRIOR APPLICATION NUMBER: US 08/9928862
/ PRIOR FILING DATE: 1997-09-12
/ PRIOR APPLICATION NUMBER: US 08/751225
/ PRIOR FILING DATE: 1996-11-15
/ PRIOR APPLICATION NUMBER: US 60/026792
/ PRIOR FILING DATE: 1996-09-27
/ NUMBER OF SEQ ID NOS: 5
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 4
/ LENGTH: 33
/ TYPE: PRT
/ ORGANISM: Homo Sapiens
US-10-858-545-4

Query Match          100.0%; Score 31; DB 5; Length 33;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRVAR 6
   |||||
Db 17 FSRVAR 22

RESULT 8
US-10-858-286-4
/ Sequence 4, Application US/10858286
/ Publication No. US20040242486A1
/ GENERAL INFORMATION:
/ APPLICANT: Chau, Raymond M.W.
/ TITLE OF INVENTION: Isolation and Use of Motoneurontropic Factors
/ FILE REFERENCE: 12592-3
/ CURRENT APPLICATION NUMBER: US/10/858,286
/ CURRENT FILING DATE: 2004-06-01
/ PRIOR APPLICATION NUMBER: US/09/592,018
/ PRIOR FILING DATE: 2000-06-12
```

```
/ PRIOR APPLICATION NUMBER: US 08/9928862
/ PRIOR FILING DATE: 1997-09-12
/ PRIOR APPLICATION NUMBER: US 08/751225
/ PRIOR FILING DATE: 1996-11-15
/ PRIOR APPLICATION NUMBER: US 60/026792
/ PRIOR FILING DATE: 1996-09-27
/ NUMBER OF SEQ ID NOS: 5
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 4
/ LENGTH: 33
/ TYPE: PRT
/ ORGANISM: Homo Sapiens
US-10-858-286-4

Query Match          100.0%; Score 31; DB 5; Length 33;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRVAR 6
   |||||
Db 17 FSRVAR 22

RESULT 9
US-10-858-543-4
/ Sequence 4, Application US/10858543
/ Publication No. US2004026691A1
/ GENERAL INFORMATION:
/ APPLICANT: Chau, Raymond M.W.
/ TITLE OF INVENTION: Isolation and Use of Motoneurontropic Factors
/ FILE REFERENCE: 12592-3
/ CURRENT APPLICATION NUMBER: US/10/858,543
/ CURRENT FILING DATE: 2004-06-01
/ PRIOR APPLICATION NUMBER: US/09/592,018
/ PRIOR FILING DATE: 2000-06-12
/ PRIOR APPLICATION NUMBER: US 08/9928862
/ PRIOR FILING DATE: 1997-09-12
/ PRIOR APPLICATION NUMBER: US 08/751225
/ PRIOR FILING DATE: 1996-11-15
/ PRIOR APPLICATION NUMBER: US 60/026792
/ PRIOR FILING DATE: 1996-09-27
/ NUMBER OF SEQ ID NOS: 5
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 4
/ LENGTH: 33
/ TYPE: PRT
/ ORGANISM: Homo Sapiens
US-10-858-543-4

Query Match          100.0%; Score 31; DB 5; Length 33;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRVAR 6
   |||||
Db 17 FSRVAR 22

RESULT 10
US-10-858-144-4
/ Sequence 4, Application US/10858144
/ Publication No. US2006002565A1
/ GENERAL INFORMATION:
/ APPLICANT: Chau, Raymond M.W.
/ TITLE OF INVENTION: Isolation and Use of Motoneurontropic Factors
/ FILE REFERENCE: 12592-3
/ CURRENT APPLICATION NUMBER: US/10/858,144
/ CURRENT FILING DATE: 2004-06-01
/ PRIOR APPLICATION NUMBER: US/09/592,018
/ PRIOR FILING DATE: 2000-06-12
/ PRIOR APPLICATION NUMBER: US 08/9928862
/ PRIOR FILING DATE: 1997-09-12
/ PRIOR APPLICATION NUMBER: US 08/751225
```

PRIOR FILING DATE: 1996-11-15
PRIOR APPLICATION NUMBER: US 60/026792
PRIOR FILING DATE: 1996-09-27
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn version 3.0
SEQ ID NO 4
LENGTH: 33
TYPE: PRT
ORGANISM: Homo Sapiens
US-10-858-144-4

Query Match 100.0%; Score 31; DB 5; Length 33;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRVAR 6
Db 17 FSRVAR 22

RESULT 11
US-10-541-343-1
Sequence 1, Application US/10541343
Publication No. US20060052299A1
GENERAL INFORMATION:
APPLICANT: Chau, Raymond Ming Mah
APPLICANT: Ko, Pui-Yuk Dorothy
TITLE OF INVENTION: NMTF Peptides and Compositions and Methods of Use
FILE REFERENCE: 2001-103US
CURRENT APPLICATION NUMBER: US/10/541,343
CURRENT FILING DATE: 2005-07-05
PRIOR APPLICATION NUMBER: PCT/US2004/001468
PRIOR FILING DATE: 2004-01-21
PRIOR APPLICATION NUMBER: 60/441,722
PRIOR FILING DATE: 2003-01-21
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn version 3.3
SEQ ID NO 1
LENGTH: 33
TYPE: PRT
ORGANISM: Artificial
FEATURE:
OTHER INFORMATION: Synthetic peptide
US-10-541-343-1

Query Match 100.0%; Score 31; DB 5; Length 33;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRVAR 6
Db 17 FSRVAR 22

RESULT 12
US-11-188-298-11746
Sequence 11746, Application US/11188298
Publication No. US20060075522A1
GENERAL INFORMATION:
APPLICANT: Abad, Mark S. et al.
TITLE OF INVENTION: GENES AND USFS FOR PLANT IMPROVEMENT
FILE REFERENCE: 38-21(53452)B
CURRENT APPLICATION NUMBER: US/11/188,298
CURRENT FILING DATE: 2005-07-22
PRIOR APPLICATION NUMBER: 60/592,978
PRIOR FILING DATE: 2004-07-31
NUMBER OF SEQ ID NOS: 22569
SEQ ID NO 11746
LENGTH: 125
TYPE: PRT
ORGANISM: Desulfotolacterium hafniense
US-11-188-298-11746

Query Match 100.0%; Score 31; DB 6; Length 125;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRVAR 6
Db 76 FSRVAR 81

RESULT 13
US-10-312-273-107
Sequence 107, Application US/10312273
Publication No. US20040005667A1
GENERAL INFORMATION:
APPLICANT: CHIRON SpA
TITLE OF INVENTION: IMMUNISATION AGAINST CHLAMYDIA PNEUMONIAE
FILE REFERENCE: P025035WO
CURRENT APPLICATION NUMBER: US/10/312,273
CURRENT FILING DATE: 2002-12-20
PRIOR APPLICATION NUMBER: 0016363.4
PRIOR FILING DATE: 2000-07-03
PRIOR APPLICATION NUMBER: 0017047.2
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 0017983.8
PRIOR FILING DATE: 2000-07-21
PRIOR APPLICATION NUMBER: 0019368.0
PRIOR FILING DATE: 2000-08-07
PRIOR APPLICATION NUMBER: 0020440.4
PRIOR FILING DATE: 2000-08-18
PRIOR APPLICATION NUMBER: 0022583.9
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 0027549.5
PRIOR FILING DATE: 2000-11-10
PRIOR APPLICATION NUMBER: 0031706.5
PRIOR FILING DATE: 2000-12-22
NUMBER OF SEQ ID NOS: 664
SOFTWARE: SeqWin99, version 1.02
SEQ ID NO 107
LENGTH: 357
TYPE: PRT
ORGANISM: Chlamydia pneumoniae
US-10-312-273-107

Query Match 100.0%; Score 31; DB 4; Length 357;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRVAR 6
Db 135 FSRVAR 140

RESULT 14
US-10-289-762-977
Sequence 977, Application US/10289762
Publication No. US20040006218A1
GENERAL INFORMATION:
APPLICANT: Griffiths, R.
TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevent
FILE REFERENCE: 9710-003-999
CURRENT APPLICATION NUMBER: US/10/289,762
CURRENT FILING DATE: 2003-03-27
NUMBER OF SEQ ID NOS: 6849
SEQ ID NO 977
LENGTH: 357
TYPE: PRT
ORGANISM: Chlamydia pneumoniae
US-10-289-762-977

Query Match 100.0%; Score 31; DB 4; Length 357;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSRVAR 6
 |||||
 Db 135 FSRVAR 140

RESULT 15
 US-10-282-122A-54956
 ; Sequence 54956, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari
 ; APPLICANT: Zyckind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.
 ; APPLICANT: Xu, H.
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: ELITRA.034A
 ; CURRENT APPLICATION NUMBER: US/10/282,122A
 ; CURRENT FILING DATE: 2003-02-20
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/230,335
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR APPLICATION NUMBER: 60/230,347
 ; PRIOR FILING DATE: 2000-09-09
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/267,636
 ; PRIOR FILING DATE: 2001-02-09
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 78614
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 54956
 ; LENGTH: 357
 ; TYPE: PRT
 ; ORGANISM: Chlamydia pneumoniae
 ; US-10-282-122A-54956

Query Match 100.0%; Score 31; DB 4; Length 357;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSRVAR 6
 |||||
 Db 135 FSRVAR 140

RESULT 16
 US-10-437-963-176098
 ; Sequence 176098, Application US/10437963
 ; Publication No. US20040123343A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wu, Wei
 ; APPLICANT: Boukharov, Andrey A.
 ; APPLICANT: Barbazuk, Brad
 ; APPLICANT: Li, Ping
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53221)B
 ; CURRENT APPLICATION NUMBER: US/10/437,963
 ; CURRENT FILING DATE: 2003-05-14
 ; NUMBER OF SEQ ID NOS: 204966
 ; SEQ ID NO 176098
 ; LENGTH: 412
 ; TYPE: PRT
 ; ORGANISM: Oryza sativa
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT4530_73880C.1.pep
 ; US-10-437-963-176098

Query Match 100.0%; Score 31; DB 4; Length 412;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSRVAR 6
 |||||
 Db 117 FSRVAR 122

RESULT 17
 US-10-369-493-23258
 ; Sequence 23258, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 23258
 ; LENGTH: 463
 ; TYPE: PRT
 ; ORGANISM: Bacillus subtilis
 ; US-10-369-493-23258

Query Match 100.0%; Score 31; DB 4; Length 463;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSRVAR 6
 |||||
 Db 76 FSRVAR 81

RESULT 18
 US-11-087-099-1759
 ; Sequence 1759, Application US/11087099
 ; Publication No. US20060041961A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Abad, Mark S. et al.
 ; TITLE OF INVENTION: Genes and Uses for Plant Improvement
 ; FILE REFERENCE: 38-21(53450)B EP
 ; CURRENT APPLICATION NUMBER: US/11/087,099
 ; CURRENT FILING DATE: 2005-03-22
 ; NUMBER OF SEQ ID NOS: 12464

SEQ ID NO 1759
LENGTH: 463
TYPE: PRT
ORGANISM: Bacillus subtilis subsp. subtilis str. 168
US-11-087-099-1759

Query Match 100.0%; Score 31; DB 6; Length 463;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRYAR 6
|||||
Db 76 FSRYAR 81

RESULT 19
US-11-188-298-12707
Sequence 12707, Application US/11188298
Publication No. US20060075522A1
GENERAL INFORMATION:
APPLICANT: Abad, Mark S. et al.
TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
FILE REFERENCE: 38-21(53452)B
CURRENT APPLICATION NUMBER: US/11/188,298
CURRENT FILING DATE: 2005-07-22
PRIOR APPLICATION NUMBER: 60/592,978
PRIOR FILING DATE: 2004-07-31
NUMBER OF SEQ ID NOS: 22569
SEQ ID NO 12707
LENGTH: 463
TYPE: PRT
ORGANISM: Bacillus subtilis subsp. subtilis str. 168
US-11-188-298-12707

Query Match 100.0%; Score 31; DB 6; Length 463;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRYAR 6
|||||
Db 76 FSRYAR 81

RESULT 20
US-09-815-242-10270
Sequence 10270, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10270
LENGTH: 474
TYPE: PRT
ORGANISM: Escherichia coli
US-09-815-242-10270

Query Match 100.0%; Score 31; DB 3; Length 474;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRYAR 6
|||||
Db 157 FSRYAR 162

RESULT 21
US-10-369-493-23518
Sequence 23518, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 23518
LENGTH: 474
TYPE: PRT
ORGANISM: Escherichia coli
US-10-369-493-23518

Query Match 100.0%; Score 31; DB 4; Length 474;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRYAR 6
|||||
Db 157 FSRYAR 162

RESULT 22
US-10-450-763-54803
Sequence 54803, Application US/10450763
Publication No. US20050196754A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 790CIP3/US
CURRENT APPLICATION NUMBER: US/10/450,763
CURRENT FILING DATE: 2003-06-11
PRIOR APPLICATION NUMBER: PCT/US01/08631
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 09/540,217
PRIOR FILING DATE: 2000-03-31
PRIOR APPLICATION NUMBER: 09/649,167
PRIOR FILING DATE: 2000-08-23
NUMBER OF SEQ ID NOS: 60736
SOFTWARE: Custom
SEQ ID NO 54803
LENGTH: 1225
TYPE: PRT

```

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (72)..(120)
; OTHER INFORMATION: Tetrahydrofolate dehydrogenase/cyclohydrolase proteins domain
; OTHER INFORMATION: Identified by eMATRIX, accession number BL00766C, p-value=1.000e-
; OTHER INFORMATION: 40, raw score of 25.86
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (7)..(211)
; OTHER INFORMATION: Tetrahydrofolate dehydrogenase/cyclohyd domain identified by
; OTHER INFORMATION: Pfam, accession name THF_DHG_CYH, E-value=1.1e-134, Pfam score of
; OTHER INFORMATION: 460.9
US-10-450-763-54803

Query Match          100.0%; Score 31; DB 5; Length 1225;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 FSRVAR 6
        |||||
        |||||
Db       613 FSRVAR 618

RESULT 23
US-10-450-763-55178
; Sequence 55178, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 55178
; LENGTH: 1258
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (332)..(380)
; OTHER INFORMATION: Tetrahydrofolate dehydrogenase/cyclohydrolase proteins domain
; OTHER INFORMATION: Identified by eMATRIX, accession number BL00766C, p-value=1.000e-
; OTHER INFORMATION: 40, raw score of 25.86
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (267)..(471)
; OTHER INFORMATION: Tetrahydrofolate dehydrogenase/cyclohyd domain identified by
; OTHER INFORMATION: Pfam, accession name THF_DHG_CYH, E-value=2.9e-134, Pfam score of
; OTHER INFORMATION: 459.5
US-10-450-763-55178

Query Match          100.0%; Score 31; DB 5; Length 1258;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 FSRVAR 6
        |||||
        |||||
Db       1199 FSRVAR 1204

RESULT 24
US-10-424-599-186222
; Sequence 186222, Application US/10424599
; Publication No. US20040031072A1
```

```

; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 186222
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(67)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_139171C.1.pep
US-10-424-599-186222

Query Match          90.3%; Score 28; DB 4; Length 67;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 FSRVAR 6
        |||||
        |||||
Db       41 FSRVAR 46

RESULT 25
US-10-424-599-146582
; Sequence 146582, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 146582
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_103383C.1.pep
US-10-424-599-146582

Query Match          90.3%; Score 28; DB 4; Length 79;
Best Local Similarity 83.3%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 FSRVAR 6
        |||||
        |||||
Db       17 FSRVAR 22

RESULT 26
US-10-424-599-265564
; Sequence 265564, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
```



```

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 265564
; LENGTH: 85
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(85)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_81824C.1.pep
US-10-424-599-265564

Query Match          90.3%; Score 28; DB 4; Length 85;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 FSRYAR 6
        |||
        |||
Db      3 FSKYAR 8

RESULT 27
US-10-183-687-178
; Sequence 178, Application US/10183687
; Publication No. US20030204870A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Allen, William B.
; APPLICANT: Cahoon, Rebecca
; APPLICANT: Epelbaum, Sabine
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Harvell, Leslie T.
; APPLICANT: Kinney, Tony
; APPLICANT: Klein, Ted
; APPLICANT: Li, Changjiang
; APPLICANT: Oliveira, Igor Cunha
; APPLICANT: Sakai, Hajime
; APPLICANT: Shen, Bo
; APPLICANT: Tarczyński, Mitchell C.
; TITLE OF INVENTION: Alteration Of Oil Traits In Plants
; FILE REFERENCE: B81458 US NA
; CURRENT APPLICATION NUMBER: US/10/183,687
; CURRENT FILING DATE: 2002-06-27
; PRIOR APPLICATION NUMBER: 60/301,913
; PRIOR FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 178
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (80)
; OTHER INFORMATION: Xaa = any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (85)
; OTHER INFORMATION: Xaa = any amino acid
US-10-183-687-178

Query Match          90.3%; Score 28; DB 4; Length 90;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 FSRYAR 6

```

```

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 219020
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_131339C.1.pep
US-10-425-115-219020

Query Match          90.3%; Score 28; DB 4; Length 91;
Best Local Similarity 83.3%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 FSRYAR 6
        |||
        |||
Db      72 FSRYSR 77

RESULT 29
US-10-767-701-51064
; Sequence 51064, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yinhua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 51064
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3480-055-P1-K1-H11.pep
US-10-767-701-51064

Query Match          90.3%; Score 28; DB 4; Length 123;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 FSRYAR 6
        |||
        |||
Db      3 FSRYAK 8

RESULT 30
US-10-767-701-43382
; Sequence 43382, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.

```

```
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(5355)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; NUMBER OF SEQ ID NOS: 2004-01-29
; SEQ ID NO 43382
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SOHBI-28MAY03-C25332_1.pep
US-10-767-701-43382

Query Match          90.3%; Score 28; DB 4; Length 124;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 FSRVAR 6
        ||:|||
Db      43 FSKYAR 48

RESULT 31
US-10-183-687-174
; Sequence 174, Application US/10183687
; Publication No. US20030204870A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Allen, William B.
; APPLICANT: Cahoon, Rebecca
; APPLICANT: Epebaum, Sabine
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Harwell, Leslie T.
; APPLICANT: Jones, Todd
; APPLICANT: Kinney, Tony
; APPLICANT: Klein, Ted
; APPLICANT: Li, Changjiang
; APPLICANT: Oliveira, Igor Cunha
; APPLICANT: Sakai, Hajime
; APPLICANT: Shen, Bo
; APPLICANT: Tarczynski, Mitchell C.
; TITLE OF INVENTION: Alteration Of Oil Traits In Plants
; FILE REFERENCE: B81458 US NA
; CURRENT APPLICATION NUMBER: US/10/183,687
; PRIOR FILING DATE: 2002-06-27
; PRIOR APPLICATION NUMBER: 60/301,913
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 174
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (91)
; OTHER INFORMATION: Xaa = any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (102)
; OTHER INFORMATION: Xaa = any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (116)
; OTHER INFORMATION: Xaa = any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (120)
; OTHER INFORMATION: Xaa = any amino acid
; FEATURE:
```

```
; NAME/KEY: UNSURE
; LOCATION: (122)
; OTHER INFORMATION: Xaa = any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (125)
; OTHER INFORMATION: Xaa = any amino acid
US-10-183-687-174

Query Match          90.3%; Score 28; DB 4; Length 132;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 FSRVAR 6
        ||:|||
Db      73 FSKYAR 78

RESULT 32
US-10-501-282-478
; Sequence 478, Application US/10501282
; Publication No. US20050203280A1
; GENERAL INFORMATION:
; APPLICANT: MCMICHAEL, JOHN CALHOUN
; APPLICANT: ZAGORSKY, ROBERT JOHN
; APPLICANT: RUSSELL, DAVID PARRISH
; APPLICANT: FLETCHER, LEAH DIANE
; TITLE OF INVENTION: ALLOILOCOCUS OTITIDIS OPEN READING FRAMES (ORFS) ENCODING
; FILE REFERENCE: AM100780 L2
; CURRENT APPLICATION NUMBER: US/10/501,282
; PRIOR FILING DATE: 2004-07-09
; PRIOR APPLICATION NUMBER: 60/333,777
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 60/426,742
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: PCT/US02/36123
; NUMBER OF SEQ ID NOS: 6653
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 478
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Alloiococcus otitidis
US-10-501-282-478

Query Match          90.3%; Score 28; DB 5; Length 145;
Best Local Similarity 83.3%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 FSRVAR 6
        ||:|||
Db      122 FSRYSR 127

RESULT 33
US-10-501-282-480
; Sequence 480, Application US/10501282
; Publication No. US20050203280A1
; GENERAL INFORMATION:
; APPLICANT: MCMICHAEL, JOHN CALHOUN
; APPLICANT: ZAGORSKY, ROBERT JOHN
; APPLICANT: RUSSELL, DAVID PARRISH
; APPLICANT: FLETCHER, LEAH DIANE
; TITLE OF INVENTION: ALLOILOCOCUS OTITIDIS OPEN READING FRAMES (ORFS) ENCODING
; FILE REFERENCE: AM100780 L2
; CURRENT APPLICATION NUMBER: US/10/501,282
; PRIOR FILING DATE: 2004-07-09
; PRIOR APPLICATION NUMBER: 60/333,777
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 60/426,742
; PRIOR FILING DATE: 2002-11-18
```

;; PRIOR APPLICATION NUMBER: PCT/US02/36123
;; PRIOR FILING DATE: 2002-11-25
;; NUMBER OF SEQ ID NOS: 663
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 480
;; LENGTH: 172
;; TYPE: PRT
;; ORGANISM: *Alloicoccus oiticidis*
US-10-501-282-480

Query Match 90.3%; Score 28; DB 5; Length 172;
Best Local Similarity 83.3%; Pred. No. 2.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FSRYSR 6
Db 149 FSRYSR 154

RESULT 34
US-09-984-245-156
; Sequence 156, Application US/09984245
; Patent No. US20020165374A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 87 Human Secreted Proteins
; FILE REFERENCE: PZ004P1
; CURRENT APPLICATION NUMBER: US/09/984,245
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 09/154,707
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: PCT/US98/05311
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: US 60/041,277
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/042,344
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,276
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,281
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/048,094
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,350
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,188
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,135
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/050,937
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,187
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,099
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,352
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,186
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,069
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,095
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,131
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,096
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,355
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,160
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,351
; PRIOR FILING DATE: 1997-05-30

;; PRIOR APPLICATION NUMBER: US 60/048,154
;; PRIOR FILING DATE: 1997-05-30
;; PRIOR APPLICATION NUMBER: US 60/054,804
;; PRIOR FILING DATE: 1997-08-05
;; PRIOR APPLICATION NUMBER: US 60/056,370
;; PRIOR FILING DATE: 1997-08-19
;; PRIOR APPLICATION NUMBER: US 60/060,862
;; PRIOR FILING DATE: 1997-10-02
;; NUMBER OF SEQ ID NOS: 343
;; SOFTWARE: PatentIn Ver. 2.0
;; SBQ ID NO 156
;; LENGTH: 174
;; TYPE: PRT
;; ORGANISM: *Homo sapiens*
US-09-984-245-156

Query Match 90.3%; Score 28; DB 3; Length 174;
Best Local Similarity 83.3%; Pred. No. 2.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FSRYSR 6
Db 69 FSRYSR 74

RESULT 35
US-09-966-262-156
; Sequence 156, Application US/09966262
; Publication No. US20030050461A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 87 Human Secreted Proteins
; FILE REFERENCE: PZ004P1
; CURRENT APPLICATION NUMBER: US/09/966,262
; CURRENT FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: US 09/154,707
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: PCT/US98/05311
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: US 60/041,277
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/042,344
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,276
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,281
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/048,094
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,350
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,188
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,135
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/050,937
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,187
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,099
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,352
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,186
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,069
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,095
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,131
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,096
; PRIOR FILING DATE: 1997-05-30

```
/ PRIOR APPLICATION NUMBER: US 60/048,355
/ PRIOR FILING DATE: 1997-05-30
/ PRIOR APPLICATION NUMBER: US 60/048,160
/ PRIOR FILING DATE: 1997-05-30
/ PRIOR APPLICATION NUMBER: US 60/048,351
/ PRIOR FILING DATE: 1997-05-30
/ PRIOR APPLICATION NUMBER: US 60/048,154
/ PRIOR FILING DATE: 1997-05-30
/ PRIOR APPLICATION NUMBER: US 60/054,804
/ PRIOR FILING DATE: 1997-08-05
/ PRIOR APPLICATION NUMBER: US 60/056,370
/ PRIOR FILING DATE: 1997-08-19
/ PRIOR APPLICATION NUMBER: US 60/060,862
/ PRIOR FILING DATE: 1997-10-02
/ NUMBER OF SEQ ID NOS: 343
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO: 156
/ LENGTH: 174
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-966-262-156
```

```
Query Match          90.3%; Score 28; DB 3; Length 174;
Best Local Similarity 83.3%; Pred. No. 2.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 FSRVAR 6
        |||||
Db       69 FARVAR 74
```

```
RESULT 36
US-09-983-966-156
/ Sequence 156, Application US/09983966
/ Publication No. US20030060619A1
/ GENERAL INFORMATION:
/ APPLICANT: Young et al.
/ TITLE OF INVENTION: 87 Human Secreted Proteins
/ FILE REFERENCE: P2004P1
/ CURRENT APPLICATION NUMBER: US/09/983,966
/ CURRENT FILING DATE: 2001-10-29
/ PRIOR APPLICATION NUMBER: 09/154,707
/ PRIOR FILING DATE: 1998-09-17
/ PRIOR APPLICATION NUMBER: PCT/US98/05311
/ PRIOR FILING DATE: 1998-03-19
/ PRIOR APPLICATION NUMBER: US 60/041,277
/ PRIOR FILING DATE: 1997-03-21
/ PRIOR APPLICATION NUMBER: US 60/042,344
/ PRIOR FILING DATE: 1997-03-21
/ PRIOR APPLICATION NUMBER: US 60/041,276
/ PRIOR FILING DATE: 1997-03-21
/ PRIOR APPLICATION NUMBER: US 60/041,281
/ PRIOR FILING DATE: 1997-03-21
/ PRIOR APPLICATION NUMBER: US 60/048,094
/ PRIOR FILING DATE: 1997-05-30
/ PRIOR APPLICATION NUMBER: US 60/048,350
/ PRIOR FILING DATE: 1997-05-30
/ PRIOR APPLICATION NUMBER: US 60/048,188
/ PRIOR FILING DATE: 1997-05-30
/ PRIOR APPLICATION NUMBER: US 60/048,135
/ PRIOR FILING DATE: 1997-05-30
/ PRIOR APPLICATION NUMBER: US 60/050,937
/ PRIOR FILING DATE: 1997-05-30
/ PRIOR APPLICATION NUMBER: US 60/048,187
/ PRIOR FILING DATE: 1997-05-30
/ PRIOR APPLICATION NUMBER: US 60/048,099
/ PRIOR FILING DATE: 1997-05-30
/ PRIOR APPLICATION NUMBER: US 60/048,352
/ PRIOR FILING DATE: 1997-05-30
/ PRIOR APPLICATION NUMBER: US 60/048,186
/ PRIOR FILING DATE: 1997-05-30
/ PRIOR APPLICATION NUMBER: US 60/048,069
/ PRIOR FILING DATE: 1997-05-30
```

```
/ PRIOR APPLICATION NUMBER: US 60/048,095
/ PRIOR FILING DATE: 1997-05-30
/ PRIOR APPLICATION NUMBER: US 60/048,131
/ PRIOR FILING DATE: 1997-05-30
/ PRIOR APPLICATION NUMBER: US 60/048,096
/ PRIOR FILING DATE: 1997-05-30
/ PRIOR APPLICATION NUMBER: US 60/048,355
/ PRIOR FILING DATE: 1997-05-30
/ PRIOR APPLICATION NUMBER: US 60/048,160
/ PRIOR FILING DATE: 1997-05-30
/ PRIOR APPLICATION NUMBER: US 60/048,351
/ PRIOR FILING DATE: 1997-05-30
/ PRIOR APPLICATION NUMBER: US 60/048,154
/ PRIOR FILING DATE: 1997-05-30
/ PRIOR APPLICATION NUMBER: US 60/054,804
/ PRIOR FILING DATE: 1997-08-05
/ PRIOR APPLICATION NUMBER: US 60/056,370
/ PRIOR FILING DATE: 1997-08-19
/ PRIOR APPLICATION NUMBER: US 60/060,862
/ PRIOR FILING DATE: 1997-10-02
/ NUMBER OF SEQ ID NOS: 343
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO: 156
/ LENGTH: 174
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-983-966-156
```

```
Query Match          90.3%; Score 28; DB 3; Length 174;
Best Local Similarity 83.3%; Pred. No. 2.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 FSRVAR 6
        |||||
Db       69 FARVAR 74
```

```
RESULT 37
US-10-059-395-156
/ Sequence 156, Application US/10059395
/ Publication No. US20030018180A1
/ GENERAL INFORMATION:
/ APPLICANT: Young et al.
/ TITLE OF INVENTION: 87 Human Secreted Proteins
/ FILE REFERENCE: P2004P1
/ CURRENT APPLICATION NUMBER: US/10/059,395
/ CURRENT FILING DATE: 2002-01-31
/ PRIOR APPLICATION NUMBER: US/09/966,262
/ PRIOR FILING DATE: 2001-10-01
/ PRIOR APPLICATION NUMBER: US 09/154,707
/ PRIOR FILING DATE: 1998-09-17
/ PRIOR APPLICATION NUMBER: PCT/US98/05311
/ PRIOR FILING DATE: 1998-03-19
/ PRIOR APPLICATION NUMBER: US 60/041,277
/ PRIOR FILING DATE: 1997-03-21
/ PRIOR APPLICATION NUMBER: US 60/042,344
/ PRIOR FILING DATE: 1997-03-21
/ PRIOR APPLICATION NUMBER: US 60/041,276
/ PRIOR FILING DATE: 1997-03-21
/ PRIOR APPLICATION NUMBER: US 60/048,094
/ PRIOR FILING DATE: 1997-05-30
/ PRIOR APPLICATION NUMBER: US 60/048,350
/ PRIOR FILING DATE: 1997-05-30
/ PRIOR APPLICATION NUMBER: US 60/048,188
/ PRIOR FILING DATE: 1997-05-30
/ PRIOR APPLICATION NUMBER: US 60/048,135
/ PRIOR FILING DATE: 1997-05-30
/ PRIOR APPLICATION NUMBER: US 60/050,937
/ PRIOR FILING DATE: 1997-05-30
/ PRIOR APPLICATION NUMBER: US 60/048,187
/ PRIOR FILING DATE: 1997-05-30
```

```
; PRIOR APPLICATION NUMBER: US 60/048,099
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,352
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,186
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,069
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,095
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,131
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,096
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,355
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,160
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,351
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,154
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/054,804
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: US 60/056,370
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: US 60/060,862
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 343
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 156
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-059-195-156
```

```
Query Match          90.3%; Score 28; DB 4; Length 174;
Best Local Similarity 83.3%; Pred. No. 2.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 FSRYAR 6
        |:||||
Db      69 FARYAR 74
```

```
RESULT 38
US-10-143-090-156
; Sequence 156, Application US/10143090
; Publication No. US20030069406A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 87 Human Secreted Proteins
; FILE REFERENCE: P2004P1
; CURRENT APPLICATION NUMBER: US/10/143,090
; CURRENT FILING DATE: 2002-05-13
; PRIOR APPLICATION NUMBER: 09/154,707
; PRIOR FILING DATE: 1998-09-17
; NUMBER OF SEQ ID NOS: 343
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 156
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-143-090-156
```

```
Query Match          90.3%; Score 28; DB 4; Length 174;
Best Local Similarity 83.3%; Pred. No. 2.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 FSRYAR 6
        |:||||
Db      69 FARYAR 74
```

```
RESULT 39
US-10-960-251-156
; Sequence 156, Application US/10960251
; Publication No. US20050089911A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 87 Human Secreted Proteins
; FILE REFERENCE: P2004P1C3
; CURRENT APPLICATION NUMBER: US/10/960,251
; CURRENT FILING DATE: 2004-10-08
; PRIOR APPLICATION NUMBER: US 10/143,090
; PRIOR FILING DATE: 2002-05-13
; PRIOR APPLICATION NUMBER: US 09/154,707
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: PCT/US98/05311
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/041,277
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: 60/042,344
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: 60/041,276
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: 60/041,281
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: 60/048,094
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,350
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,188
; PRIOR FILING DATE: 1997-05-30
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 343
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 156
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-960-251-156
```

```
Query Match          90.3%; Score 28; DB 5; Length 174;
Best Local Similarity 83.3%; Pred. No. 2.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 FSRYAR 6
        |:||||
Db      69 FARYAR 74
```

```
RESULT 40
US-10-437-963-197182
; Sequence 197182, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 197182
; LENGTH: 179
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
```

OTHER INFORMATION: Clone ID: PAT_MRT4530_92964C.1.pep
US-10-437-963-197182

Query Match 90.3%; Score 28; DB 4; Length 179;
Best Local Similarity 83.3%; Pred. No. 3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRVAR 6
||:||||
DB 10 FARVAR 15

RESULT 41
US-11-096-568A-9600

Sequence 9600, Application US/11096568A
Publication No. US20060048240A1

GENERAL INFORMATION:

APPLICANT: Alexandrov, Nikolai et al.

TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides

FILE REFERENCE: 2750-1592PUS2

CURRENT APPLICATION NUMBER: US/11/096,568A

CURRENT FILING DATE: 2005-04-01

NUMBER OF SEQ ID NOS: 34471

SEQ ID NO 9600

LENGTH: 225

TYPE: PRT

ORGANISM: Triticum aestivum

FEATURE:

NAME/KEY: misc_feature

LOCATION: (1)..(225)

OTHER INFORMATION: Ceres Seq. ID no. 12614812

US-11-096-568A-9600

Query Match 90.3%; Score 28; DB 6; Length 225;
Best Local Similarity 83.3%; Pred. No. 3.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRVAR 6
||:||||
DB 144 FSKVAR 149

RESULT 42
US-10-425-115-314441

Sequence 314441, Application US/10425115
Publication No. US20040214272A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

FILE REFERENCE: 38-21(53222)B

CURRENT APPLICATION NUMBER: US/10/425,115

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 369326

SEQ ID NO 314441

LENGTH: 234

TYPE: PRT

ORGANISM: Zea mays

FEATURE:

OTHER INFORMATION: Clone ID: MRT4577_4982C.1.pep

US-10-425-115-314441

Query Match 90.3%; Score 28; DB 4; Length 234;
Best Local Similarity 83.3%; Pred. No. 3.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRVAR 6
||:||||
DB 153 FSKVAR 158

RESULT 43
US-10-424-599-254640

Sequence 254640, Application US/10424599
Publication No. US20040031072A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

FILE REFERENCE: 38-21(53223)B

CURRENT APPLICATION NUMBER: US/10/424,599

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 285684

SEQ ID NO 254640

LENGTH: 239

TYPE: PRT

ORGANISM: Glycine max

FEATURE:

OTHER INFORMATION: Clone ID: PAT_MRT3847_71963C.1.pep

US-10-424-599-254640

Query Match 90.3%; Score 28; DB 4; Length 239;
Best Local Similarity 83.3%; Pred. No. 4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRVAR 6
||:||||
DB 157 FSKVAR 162

RESULT 44
US-10-987-454-122

Sequence 122, Application US/10987454
Publication No. US20050172359A1

GENERAL INFORMATION:

APPLICANT: Moloney, Maurice

APPLICANT: Reid, Alexandra

TITLE OF INVENTION: METHODS FOR THE PRODUCTION OF APOLIPOPROTEINS IN TRANSGENIC

FILE REFERENCE: 9369-311

CURRENT APPLICATION NUMBER: US/10/987,454

CURRENT FILING DATE: 2004-11-15

PRIOR APPLICATION NUMBER: US 60/519,606

PRIOR FILING DATE: 2003-11-14

PRIOR APPLICATION NUMBER: US 60/579,733

PRIOR FILING DATE: 2004-06-16

NUMBER OF SEQ ID NOS: 251

SOFTWARE: PatentIn version 3.3

SEQ ID NO 122

LENGTH: 243

TYPE: PRT

ORGANISM: Arabidopsis thaliana

US-10-987-454-122

Query Match 90.3%; Score 28; DB 5; Length 243;
Best Local Similarity 83.3%; Pred. No. 4.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRVAR 6
||:||||
DB 159 FSKVAR 164

RESULT 45
US-10-183-687-160

Sequence 160, Application US/10183687
Publication No. US20030204870A1

GENERAL INFORMATION:

APPLICANT: Allen, Steve

```
; APPLICANT: Allen, William B.
; APPLICANT: Cahoon, Rebecca
; APPLICANT: Epelbaum, Sabine
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Harvell, Leslie T.
; APPLICANT: Jones, Todd
; APPLICANT: Kinney, Tony
; APPLICANT: Klein, Ted
; APPLICANT: Li, Changjiang
; APPLICANT: Oliveira, Igor Cunha
; APPLICANT: Sakai, Hajime
; APPLICANT: Shen, Bo
; APPLICANT: Tarczyński, Mitchell C.
; TITLE OF INVENTION: Alteration Of Oil Traits In Plants
; FILE REFERENCE: BBI458 US NA
; CURRENT APPLICATION NUMBER: US/10/183,687
; CURRENT FILING DATE: 2002-06-27
; PRIOR APPLICATION NUMBER: 60/301,913
; PRIOR FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 160
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Zea mays
US-10-183-687-160
```

```
Query Match          90.3% Score 28; DB 4; Length 244;
Best Local Similarity 83.3% Pred. No. 4.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 FSRVAR 6
        ||:||||
Db      163 FSKYAR 168
```

```
RESULT 46
US-10-425-114-54567
; Sequence 54567, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 54567
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700838356_Flt.pep
US-10-425-114-54567
```

```
Query Match          90.3% Score 28; DB 4; Length 256;
Best Local Similarity 83.3% Pred. No. 4.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 FSRVAR 6
        ||:||||
Db      174 FSKYAR 179
```

```
RESULT 47
US-10-437-963-133297
; Sequence 133297, Application US/10437963
```

```
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 133297
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_35181C.1.pep
US-10-437-963-133297
```

```
Query Match          90.3% Score 28; DB 4; Length 285;
Best Local Similarity 83.3% Pred. No. 4.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 FSRVAR 6
        ||:||||
Db      204 FSKYAR 209
```

```
RESULT 48
US-10-987-454-127
; Sequence 127, Application US/10987454
; Publication No. US20050172359A1
; GENERAL INFORMATION:
; APPLICANT: Reid, Alexandra
; APPLICANT: Moloney, Maurice
; TITLE OF INVENTION: METHODS FOR THE PRODUCTION OF APOLIPOPROTEINS IN TRANSGENIC
; FILE REFERENCE: 9369-311
; CURRENT APPLICATION NUMBER: US/10/987,454
; CURRENT FILING DATE: 2004-11-15
; PRIOR APPLICATION NUMBER: US 60/519,606
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: US 60/579,733
; PRIOR FILING DATE: 2004-06-16
; NUMBER OF SEQ ID NOS: 251
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 127
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Hordeum vulgare
US-10-987-454-127
```

```
Query Match          90.3% Score 28; DB 5; Length 301;
Best Local Similarity 83.3% Pred. No. 5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 FSRVAR 6
        ||:||||
Db      220 FSKYAR 225
```

```
RESULT 49
US-10-987-454-128
; Sequence 128, Application US/10987454
; Publication No. US20050172359A1
; GENERAL INFORMATION:
; APPLICANT: Reid, Alexandra
; APPLICANT: Moloney, Maurice
```

```

; TITLE OF INVENTION: METHODS FOR THE PRODUCTION OF APOLIPOPROTEINS IN TRANSGENIC
; TITLE OF INVENTION: PLANTS
; FILE REFERENCE: 9369-311
; CURRENT APPLICATION NUMBER: US/10/987,454
; CURRENT FILING DATE: 2004-11-15
; PRIOR APPLICATION NUMBER: US 60/519,606
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: US 60/579,733
; PRIOR FILING DATE: 2004-06-16
; NUMBER OF SEQ ID NOS: 251
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 128
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Hordeum vulgare
US-10-987-454-128

```

```

Query Match          90.3%; Score 28; DB 5; Length 301;
Best Local Similarity 83.3%; Pred. No. 5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 FSRVAR 6
        ||:||||
Db      220 FSKVAR 225

```

```

RESULT 50
US-10-450-763-37522
; Sequence 37522, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 37522
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(301)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-450-763-37522

```

```

Query Match          90.3%; Score 28; DB 5; Length 301;
Best Local Similarity 83.3%; Pred. No. 5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 FSRVAR 6
        ||:||||
Db      90 FARVAR 95

```

```

Search completed: August 29, 2006, 10:18:53
Job time : 75.5941 secs

```


GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2006, 09:43:08 ; Search time 4.75248 Seconds
(without alignments)
86.383 Million cell updates/sec

Title: US-10-541-343-2

Perfect score: 31

Sequence: 1 FSRXAR 6

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 247503 seqs, 68422524 residues

Total number of hits satisfying chosen parameters: 247503

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database :

Published Applications AA New:*

- 1: /EMC_Celerra_SIDS3/ptodata/2/pubppaa/US09_NEW_PUB pep:*
- 2: /EMC_Celerra_SIDS3/ptodata/2/pubppaa/US06_NEW_PUB pep:*
- 3: /EMC_Celerra_SIDS3/ptodata/2/pubppaa/US07_NEW_PUB pep:*
- 4: /EMC_Celerra_SIDS3/ptodata/2/pubppaa/US08_NEW_PUB pep:*
- 5: /EMC_Celerra_SIDS3/ptodata/2/pubppaa/PCR_NEW_PUB pep:*
- 6: /EMC_Celerra_SIDS3/ptodata/2/pubppaa/US10_NEW_PUB pep:*
- 7: /EMC_Celerra_SIDS3/ptodata/2/pubppaa/US11_NEW_PUB pep:*
- 8: /EMC_Celerra_SIDS3/ptodata/2/pubppaa/US60_NEW_PUB pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	28	90.3	174	7	US-11-105-223-53
2	28	90.3	225	7	US-11-056-355B-68811
3	28	90.3	240	7	US-11-056-355B-106575
4	28	90.3	240	7	US-11-056-355B-117814
5	28	90.3	243	7	US-11-056-355B-106574
6	28	90.3	243	7	US-11-056-355B-117813
7	28	90.3	255	7	US-11-056-355B-106573
8	28	90.3	255	7	US-11-056-355B-117812
9	28	90.3	301	7	US-11-056-355B-68810
10	28	90.3	348	6	US-10-449-902-37988
11	28	90.3	358	6	US-10-953-349-9971
12	27	87.1	129	6	US-10-953-349-7495
13	27	87.1	208	6	US-10-953-349-2963
14	27	87.1	209	6	US-10-953-349-2963
15	27	87.1	216	6	US-10-953-349-2962
16	27	87.1	222	6	US-10-953-349-7494
17	27	87.1	290	6	US-10-953-349-9130
18	27	87.1	312	7	US-11-174-307B-202
19	27	87.1	323	6	US-10-449-902-56777
20	27	87.1	334	6	US-10-953-349-21500
21	27	87.1	338	6	US-10-449-902-34829
22	27	87.1	353	6	US-10-953-349-21499
23	27	87.1	374	6	US-10-449-902-33374
24	27	87.1	374	6	US-10-449-902-45028
25	27	87.1	464	6	US-10-953-349-21498

26	27	87.1	478	7	US-11-330-403-7004	Sequence 7004, Ap
27	27	87.1	488	7	US-11-399-675-7	Sequence 7, Appli
28	27	87.1	488	7	US-11-353-390A-7	Sequence 7, Appli
29	27	87.1	488	7	US-11-354-310A-7	Sequence 7, Appli
30	27	87.1	541	7	US-11-174-307B-4272	Sequence 4272, Ap
31	26	83.9	107	6	US-10-449-902-33781	Sequence 33781, A
32	26	83.9	109	7	US-11-056-355B-581	Sequence 581, App
33	26	83.9	129	7	US-11-337-300-451	Sequence 451, App
34	26	83.9	217	6	US-10-449-902-31883	Sequence 31883, App
35	26	83.9	219	7	US-11-056-355B-16531	Sequence 16531, A
36	26	83.9	256	7	US-11-056-355B-16851	Sequence 16851, A
37	26	83.9	256	7	US-11-337-300-286	Sequence 286, App
38	26	83.9	271	6	US-10-449-902-30299	Sequence 30299, A
39	26	83.9	271	6	US-10-449-902-31725	Sequence 31725, A
40	26	83.9	271	6	US-10-449-902-32135	Sequence 32135, A
41	26	83.9	271	6	US-10-449-902-51604	Sequence 51604, A
42	26	83.9	271	7	US-11-056-355B-7374	Sequence 7374, Ap
43	26	83.9	311	7	US-11-056-355B-16849	Sequence 16849, A
44	26	83.9	311	7	US-11-056-355B-7372	Sequence 7372, Ap
45	26	83.9	334	7	US-11-330-403-9151	Sequence 9151, Ap
46	26	83.9	334	7	US-11-330-403-19189	Sequence 19189, A
47	26	83.9	334	7	US-11-404-745-2	Sequence 2, Appli
48	26	83.9	343	7	US-11-359-554-3	Sequence 3, Appli
49	26	83.9	348	6	US-10-471-571A-4440	Sequence 4440, Ap
50	26	83.9	365	7	US-11-333-747A-22	Sequence 22, Appl
51	26	83.9	377	7	US-11-056-355B-16529	Sequence 16529, A
52	26	83.9	415	6	US-10-953-349-27380	Sequence 27380, A
53	26	83.9	415	7	US-11-056-355B-63199	Sequence 63199, A
54	26	83.9	439	6	US-10-449-902-55850	Sequence 55850, A
55	26	83.9	454	7	US-11-330-403-1946	Sequence 1946, Ap
56	26	83.9	461	7	US-11-330-403-18068	Sequence 18068, A
57	26	83.9	463	6	US-10-471-571A-3898	Sequence 3898, Ap
58	26	83.9	474	7	US-11-330-403-13271	Sequence 13271, A
59	26	83.9	477	6	US-10-953-349-27379	Sequence 27379, A
60	26	83.9	495	6	US-11-056-355B-63198	Sequence 63198, A
61	26	83.9	495	6	US-10-953-349-27378	Sequence 27378, A
62	26	83.9	495	7	US-11-056-355B-63197	Sequence 63197, A
63	26	83.9	509	6	US-10-471-571A-3898	Sequence 3898, Ap
64	26	83.9	520	6	US-10-471-571A-5356	Sequence 5356, Ap
65	26	83.9	730	7	US-11-056-355B-101800	Sequence 101800, A
66	26	83.9	730	7	US-11-056-355B-113039	Sequence 113039, A
67	26	83.9	739	7	US-11-056-355B-101799	Sequence 101799, A
68	26	83.9	739	7	US-11-056-355B-113038	Sequence 113038, A
69	25	80.6	98	7	US-11-330-403-15890	Sequence 15890, A
70	25	80.6	207	7	US-11-330-403-577	Sequence 577, App
71	25	80.6	216	6	US-10-449-902-56102	Sequence 56102, A
72	25	80.6	218	7	US-11-056-355B-13911	Sequence 13911, A
73	25	80.6	218	7	US-11-056-355B-21841	Sequence 21841, A
74	25	80.6	219	7	US-11-056-355B-29799	Sequence 29799, A
75	25	80.6	219	7	US-11-056-355B-33389	Sequence 33389, A
76	25	80.6	219	7	US-11-056-355B-78387	Sequence 78387, A
77	25	80.6	219	7	US-11-330-403-10769	Sequence 10769, A
78	25	80.6	235	6	US-10-449-902-32366	Sequence 32366, A
79	25	80.6	235	6	US-10-449-902-88414	Sequence 88414, A
80	25	80.6	239	7	US-11-056-355B-80982	Sequence 80982, A
81	25	80.6	245	7	US-11-056-355B-80981	Sequence 80981, A
82	25	80.6	245	7	US-11-056-355B-21840	Sequence 21840, A
83	25	80.6	249	7	US-11-056-355B-80980	Sequence 80980, A
84	25	80.6	273	7	US-11-330-403-3987	Sequence 3987, Ap
85	25	80.6	415	7	US-11-056-355B-87756	Sequence 87756, A
86	25	80.6	415	7	US-11-056-355B-87755	Sequence 87755, A
87	25	80.6	437	7	US-11-056-355B-87754	Sequence 87754, A
88	25	80.6	459	7	US-11-226-554-92	Sequence 92, Appli
89	25	80.6	459	7	US-11-226-554-92	Sequence 92, Appli
90	25	80.6	475	6	US-10-449-902-91911	Sequence 91911, A
91	25	80.6	493	6	US-10-449-902-37649	Sequence 37649, A
92	25	80.6	511	7	US-11-313-450-26	Sequence 26, Appli
93	25	80.6	528	6	US-10-449-902-44124	Sequence 44124, A
94	25	80.6	596	6	US-10-449-902-34689	Sequence 34689, A
95	25	80.6	604	7	US-11-056-355B-55018	Sequence 55018, A
96	25	80.6	612	6	US-10-449-902-51767	Sequence 51767, A
97	25	80.6	620	7	US-11-056-355B-38630	Sequence 38630, A
98	25	80.6	622	7	US-11-056-355B-55017	Sequence 55017, A

99	25	80.6	626	7	US-11-056-355B-46446	Sequence 46446, A	172	24	77.4	491	7	US-11-330-403-4874	Sequence 4874, Ap
100	25	80.6	626	7	US-11-056-355B-47112	Sequence 47112, A	173	24	77.4	491	7	US-11-330-403-7107	Sequence 7107, Ap
101	25	80.6	626	7	US-11-056-355B-49228	Sequence 49228, A	174	24	77.4	499	7	US-11-377-884-42	Sequence 42, Appl
102	25	80.6	656	7	US-11-056-355B-14120	Sequence 14120, A	175	24	77.4	502	7	US-11-056-355B-45190	Sequence 45190, A
103	25	80.6	675	7	US-11-056-355B-38629	Sequence 38629, A	176	24	77.4	503	6	US-10-532-868-16	Sequence 16, Appl
104	25	80.6	675	7	US-11-056-355B-55016	Sequence 55016, A	177	24	77.4	518	7	US-11-056-355B-45189	Sequence 45189, A
105	25	80.6	689	7	US-11-056-355B-38628	Sequence 38628, A	178	24	77.4	523	7	US-11-330-403-6029	Sequence 6029, Ap
106	25	80.6	693	7	US-11-056-355B-14119	Sequence 14119, A	179	24	77.4	527	6	US-10-449-902-40991	Sequence 40981, A
107	25	80.6	722	7	US-10-449-902-44559	Sequence 44559, A	180	24	77.4	537	6	US-10-449-902-42565	Sequence 42565, A
108	25	80.6	722	7	US-11-056-355B-44118	Sequence 44118, A	181	24	77.4	558	7	US-11-051-725-137	Sequence 137, App
109	25	80.6	774	7	US-11-056-355B-46445	Sequence 46445, A	182	24	77.4	561	7	US-11-330-403-18408	Sequence 18408, A
110	25	80.6	774	7	US-11-056-355B-47111	Sequence 47111, A	183	24	77.4	563	7	US-11-377-884-21	Sequence 21, Appl
111	25	80.6	774	7	US-11-056-355B-49227	Sequence 49227, A	184	24	77.4	605	6	US-10-449-902-45218	Sequence 45218, A
112	25	80.6	798	7	US-11-056-355B-46444	Sequence 46444, A	185	24	77.4	647	7	US-11-330-403-5831	Sequence 5831, Ap
113	25	80.6	798	7	US-11-056-355B-47110	Sequence 47110, A	186	24	77.4	807	7	US-11-330-403-37372	Sequence 4372, Ap
114	25	80.6	798	7	US-11-056-355B-49226	Sequence 49226, A	187	24	77.4	816	7	US-11-330-403-45498	Sequence 45498, Ap
115	25	80.6	822	7	US-11-328-747-12	Sequence 12, Appl	188	24	77.4	932	7	US-11-051-725-133	Sequence 133, App
116	25	80.6	822	7	US-11-090-563-12	Sequence 12, Appl	189	24	77.4	989	7	US-11-051-725-4	Sequence 4, Appl
117	25	80.6	923	7	US-11-330-403-4339	Sequence 4339, Ap	190	24	77.4	2105	7	US-11-051-725-9	Sequence 9, Appl
118	25	80.6	1005	6	US-10-449-902-53787	Sequence 53787, A	191	24	77.4	2123	7	US-11-051-725-180	Sequence 180, App
119	25	80.6	2397	6	US-10-471-571A-2632	Sequence 53787, A	192	24	77.4	2169	7	US-11-051-725-5	Sequence 5, Appl
120	25	80.6	3429	6	US-10-953-349-9275	Sequence 9275, Ap	193	24	77.4	2182	7	US-11-051-725-8	Sequence 8, Appl
121	25	80.6	3438	6	US-10-953-349-9273	Sequence 9273, Ap	194	24	77.4	2192	7	US-11-051-725-147	Sequence 147, App
122	25	80.6	3478	6	US-10-953-349-9273	Sequence 9273, Ap	195	24	77.4	2221	6	US-10-829-000-3	Sequence 3, Appl
123	24	77.4	71	7	US-11-056-355B-29148	Sequence 29148, A	196	24	77.4	2221	6	US-11-051-725-162	Sequence 162, App
124	24	77.4	71	7	US-11-056-355B-32738	Sequence 32738, A	197	23	74.2	9	7	US-11-342-208A-6	Sequence 6, Appl
125	24	77.4	177	6	US-10-953-349-33412	Sequence 33412, A	198	23	74.2	10	6	US-10-538-066-73	Sequence 73, Appl
126	24	77.4	177	6	US-10-953-349-33412	Sequence 33412, A	199	23	74.2	10	6	US-10-544-050-37	Sequence 37, Appl
127	24	77.4	201	6	US-10-953-349-14117	Sequence 14117, A	200	23	74.2	10	6	US-10-544-050-38	Sequence 38, Appl
128	24	77.4	201	6	US-10-449-902-30457	Sequence 30457, A	201	23	74.2	13	6	US-10-538-066-443	Sequence 443, App
129	24	77.4	201	6	US-10-449-902-30457	Sequence 30457, A	202	23	74.2	21	6	US-10-449-902-36311	Sequence 36311, A
130	24	77.4	201	6	US-11-056-355B-56385	Sequence 56385, A	203	23	74.2	44	6	US-10-449-902-54036	Sequence 54036, A
131	24	77.4	212	6	US-10-953-349-5598	Sequence 5598, Ap	204	23	74.2	59	6	US-10-449-902-28990	Sequence 3850, A
132	24	77.4	214	6	US-10-953-349-14116	Sequence 14116, A	205	23	74.2	95	6	US-10-553-436-334	Sequence 334, App
133	24	77.4	214	6	US-10-449-902-30361	Sequence 30361, A	206	23	74.2	95	7	US-11-043-882-185	Sequence 385, App
134	24	77.4	214	6	US-10-449-902-33402	Sequence 33402, A	207	23	74.2	112	6	US-10-544-050-2	Sequence 2, Appl
135	24	77.4	214	6	US-10-449-902-55847	Sequence 55847, A	208	23	74.2	112	6	US-10-544-050-5	Sequence 5, Appl
136	24	77.4	221	6	US-11-056-355B-56384	Sequence 56384, A	209	23	74.2	112	6	US-10-544-050-6	Sequence 6, Appl
137	24	77.4	228	6	US-10-953-349-33411	Sequence 33411, A	210	23	74.2	126	6	US-10-553-436-243	Sequence 243, App
138	24	77.4	228	6	US-11-056-355B-18832	Sequence 18832, A	211	23	74.2	131	7	US-11-304-986-4	Sequence 4, Appl
139	24	77.4	237	7	US-10-953-349-14115	Sequence 14115, A	212	23	74.2	131	7	US-11-304-986-6	Sequence 6, Appl
140	24	77.4	237	7	US-11-056-355B-56383	Sequence 56383, A	213	23	74.2	133	6	US-10-553-436-242	Sequence 242, App
141	24	77.4	240	6	US-10-953-349-5596	Sequence 5596, Ap	214	23	74.2	133	7	US-11-330-403-6534	Sequence 6534, Ap
142	24	77.4	243	6	US-10-953-349-33410	Sequence 33410, A	215	23	74.2	141	6	US-10-449-902-35544	Sequence 35544, A
143	24	77.4	243	6	US-11-056-355B-18831	Sequence 18831, A	216	23	74.2	141	7	US-11-056-355B-51908	Sequence 51908, A
144	24	77.4	285	6	US-10-449-902-49541	Sequence 49541, A	217	23	74.2	144	7	US-11-330-403-7866	Sequence 7866, Ap
145	24	77.4	286	6	US-10-374-780A-1769	Sequence 1769, Ap	218	23	74.2	145	6	US-10-953-349-52378	Sequence 52378, A
146	24	77.4	353	6	US-10-953-349-20094	Sequence 20094, A	219	23	74.2	145	6	US-11-056-355B-52378	Sequence 52378, A
147	24	77.4	353	6	US-11-056-355B-60417	Sequence 60417, A	220	23	74.2	148	6	US-10-953-349-18111	Sequence 18111, A
148	24	77.4	357	7	US-11-056-355B-11256	Sequence 11256, A	221	23	74.2	150	6	US-10-553-436-244	Sequence 244, App
149	24	77.4	379	6	US-10-449-902-36694	Sequence 36694, A	222	23	74.2	152	7	US-11-056-355B-29423	Sequence 29423, A
150	24	77.4	379	6	US-10-449-902-50338	Sequence 50338, A	223	23	74.2	152	7	US-11-056-355B-33013	Sequence 33013, A
151	24	77.4	379	7	US-11-056-355B-11255	Sequence 11255, A	224	23	74.2	152	7	US-11-056-355B-99313	Sequence 99313, A
152	24	77.4	385	6	US-10-449-902-50074	Sequence 50074, A	225	23	74.2	152	7	US-11-056-355B-110552	Sequence 110552, A
153	24	77.4	386	7	US-11-056-355B-43250	Sequence 43250, A	226	23	74.2	153	7	US-11-056-355B-51907	Sequence 51907, A
154	24	77.4	386	7	US-11-056-355B-74637	Sequence 74637, A	227	23	74.2	154	7	US-11-056-355B-29422	Sequence 29422, A
155	24	77.4	386	7	US-11-056-355B-103331	Sequence 103331, A	228	23	74.2	154	7	US-11-056-355B-33012	Sequence 33012, A
156	24	77.4	386	7	US-11-056-355B-114570	Sequence 114570, A	229	23	74.2	154	7	US-11-056-355B-99312	Sequence 99312, A
157	24	77.4	389	7	US-11-056-355B-43249	Sequence 43249, A	230	23	74.2	154	7	US-11-056-355B-110551	Sequence 110551, A
158	24	77.4	389	7	US-11-056-355B-74636	Sequence 74636, A	231	23	74.2	155	6	US-10-553-436-241	Sequence 241, App
159	24	77.4	389	7	US-11-056-355B-103330	Sequence 103330, A	232	23	74.2	156	6	US-11-056-355B-29421	Sequence 29421, App
160	24	77.4	389	7	US-11-056-355B-114569	Sequence 114569, A	233	23	74.2	156	7	US-11-056-355B-33011	Sequence 33011, A
161	24	77.4	393	7	US-11-056-355B-103329	Sequence 103329, A	234	23	74.2	156	7	US-11-056-355B-99311	Sequence 99311, A
162	24	77.4	393	7	US-11-056-355B-114568	Sequence 114568, A	235	23	74.2	157	7	US-11-056-355B-110550	Sequence 110550, A
163	24	77.4	401	7	US-11-056-355B-11254	Sequence 11254, A	236	23	74.2	159	7	US-11-330-403-5576	Sequence 5576, Ap
164	24	77.4	405	7	US-11-056-355B-43248	Sequence 43248, A	237	23	74.2	162	6	US-10-449-902-45762	Sequence 45762, A
165	24	77.4	407	7	US-11-056-355B-74635	Sequence 74635, A	238	23	74.2	174	7	US-11-056-355B-51906	Sequence 51906, A
166	24	77.4	410	7	US-11-330-403-13856	Sequence 13856, A	239	23	74.2	178	7	US-11-056-355B-44007	Sequence 44007, A
167	24	77.4	431	6	US-10-449-902-51817	Sequence 51817, A	240	23	74.2	181	6	US-10-953-349-30541	Sequence 30541, A
168	24	77.4	439	6	US-10-449-902-51817	Sequence 64, Appl	241	23	74.2	181	7	US-11-056-355B-61218	Sequence 61218, A
169	24	77.4	465	7	US-11-051-725-129	Sequence 129, App	242	23	74.2	184	6	US-10-553-436-335	Sequence 335, App
170	24	77.4	485	7	US-11-051-725-154	Sequence 154, App	243	23	74.2	198	6	US-10-449-902-29036	Sequence 29036, A
171	24	77.4	485	7	US-11-051-725-170	Sequence 170, App	244	23	74.2	201	6	US-10-449-902-88412	Sequence 88412, A

245	23	74.2	201	7	US-11-056-355B-44006	Sequence 44006, A
246	23	74.2	217	7	US-11-056-355B-57477	Sequence 57477, A
247	23	74.2	217	7	US-11-330-403-3089	Sequence 3089, Ap
248	23	74.2	217	7	US-11-330-403-9378	Sequence 9378, Ap
249	23	74.2	230	7	US-11-056-355B-69587	Sequence 69587, A
250	23	74.2	233	7	US-11-056-355B-80960	Sequence 80960, A
251	23	74.2	236	7	US-11-056-355B-44005	Sequence 44005, A
252	23	74.2	239	6	US-10-953-349-10540	Sequence 10540, A
253	23	74.2	239	6	US-11-056-355B-61217	Sequence 61217, A
254	23	74.2	241	6	US-10-953-349-20142	Sequence 20142, A
255	23	74.2	242	7	US-11-056-355B-83932	Sequence 83932, A
256	23	74.2	244	6	US-10-953-349-30539	Sequence 30539, A
257	23	74.2	244	7	US-11-056-355B-61216	Sequence 61216, A
258	23	74.2	246	7	US-11-330-403-8177	Sequence 8177, Ap
259	23	74.2	250	6	US-10-449-902-53359	Sequence 53359, A
260	23	74.2	251	7	US-11-251-208-433	Sequence 433, App
261	23	74.2	264	6	US-10-953-349-20141	Sequence 20141, A
262	23	74.2	265	7	US-11-056-355B-62734	Sequence 62734, A
263	23	74.2	270	7	US-11-249-111-73	Sequence 73, Appl
264	23	74.2	270	7	US-11-330-403-15694	Sequence 15694, A
265	23	74.2	275	6	US-10-953-349-17802	Sequence 17802, A
266	23	74.2	275	7	US-11-056-355B-54203	Sequence 54203, A
267	23	74.2	280	7	US-11-056-355B-83931	Sequence 83931, A
268	23	74.2	282	7	US-11-056-355B-71608	Sequence 71608, A
269	23	74.2	299	6	US-10-449-902-47662	Sequence 47662, A
270	23	74.2	300	6	US-10-449-902-47936	Sequence 47936, A
271	23	74.2	300	6	US-10-449-902-53281	Sequence 53281, A
272	23	74.2	301	6	US-10-449-902-30034	Sequence 30034, A
273	23	74.2	301	7	US-11-330-403-4508	Sequence 4508, Ap
274	23	74.2	322	6	US-10-953-349-17801	Sequence 17801, A
275	23	74.2	322	7	US-11-056-355B-54202	Sequence 54202, A
276	23	74.2	324	6	US-10-449-902-30795	Sequence 30795, A
277	23	74.2	324	6	US-10-449-902-49734	Sequence 49734, A
278	23	74.2	324	7	US-11-174-307B-3318	Sequence 3318, Ap
279	23	74.2	326	7	US-11-330-403-7617	Sequence 7617, Ap
280	23	74.2	330	6	US-10-449-902-51597	Sequence 51597, A
281	23	74.2	335	6	US-10-449-902-56163	Sequence 56163, A
282	23	74.2	341	7	US-11-174-307B-3548	Sequence 3548, Ap
283	23	74.2	341	6	US-11-330-403-16154	Sequence 16154, A
284	23	74.2	343	6	US-10-953-349-17800	Sequence 17800, A
285	23	74.2	343	7	US-11-056-355B-54201	Sequence 54201, A
286	23	74.2	349	7	US-11-056-355B-57436	Sequence 57436, A
287	23	74.2	349	7	US-11-056-355B-62733	Sequence 62733, A
288	23	74.2	352	7	US-11-056-355B-8677	Sequence 8677, Ap
289	23	74.2	353	6	US-10-953-349-35828	Sequence 35828, A
290	23	74.2	354	6	US-10-449-902-35334	Sequence 35334, A
291	23	74.2	354	6	US-10-374-780A-444	Sequence 444, App
292	23	74.2	360	6	US-10-449-902-36594	Sequence 36594, A
293	23	74.2	364	6	US-10-543-046-1	Sequence 1, Appli
294	23	74.2	367	6	US-10-953-349-7459	Sequence 7459, Ap
295	23	74.2	368	7	US-11-330-403-5008	Sequence 5008, Ap
296	23	74.2	368	7	US-11-330-403-10185	Sequence 10185, A
297	23	74.2	369	7	US-11-056-355B-17607	Sequence 71607, A
298	23	74.2	371	6	US-10-953-349-20140	Sequence 20140, A
299	23	74.2	372	7	US-11-293-697-2840	Sequence 2840, Ap
300	23	74.2	375	6	US-10-449-902-33919	Sequence 33919, A
301	23	74.2	377	6	US-10-953-349-35827	Sequence 35827, A
302	23	74.2	380	6	US-10-449-902-33929	Sequence 33929, A
303	23	74.2	386	6	US-10-953-349-9752	Sequence 9752, Ap
304	23	74.2	386	7	US-11-056-355B-6674	Sequence 6674, Ap
305	23	74.2	386	7	US-11-056-355B-16815	Sequence 16815, A
306	23	74.2	391	7	US-11-223-945-38	Sequence 38, Appl
307	23	74.2	394	6	US-10-953-349-9751	Sequence 9751, Ap
308	23	74.2	395	7	US-11-056-355B-9813	Sequence 9813, Ap
309	23	74.2	396	7	US-11-330-403-10174	Sequence 10174, A
310	23	74.2	397	7	US-11-056-355B-57435	Sequence 57435, A
311	23	74.2	397	7	US-11-330-403-18032	Sequence 18032, A
312	23	74.2	405	7	US-11-056-355B-45402	Sequence 45402, A
313	23	74.2	411	6	US-10-530-187-202	Sequence 202, App
314	23	74.2	413	6	US-10-449-902-46807	Sequence 46807, A
315	23	74.2	418	6	US-10-449-902-55989	Sequence 55989, A
316	23	74.2	420	7	US-11-056-355B-9812	Sequence 9812, Ap
317	23	74.2	426	7	US-11-056-355B-80331	Sequence 80331, A
318	23	74.2	428	7	US-11-056-355B-42312	Sequence 42312, A
319	23	74.2	428	7	US-11-056-355B-89166	Sequence 89166, A
320	23	74.2	428	7	US-11-056-355B-92922	Sequence 92922, A
321	23	74.2	430	6	US-10-449-902-39680	Sequence 39680, A
322	23	74.2	431	6	US-10-953-349-74336	Sequence 74336, A
323	23	74.2	431	7	US-11-056-355B-17105	Sequence 17105, A
324	23	74.2	434	7	US-11-056-355B-4946	Sequence 4946, Ap
325	23	74.2	438	7	US-11-056-355B-42311	Sequence 42311, A
326	23	74.2	438	7	US-11-056-355B-89165	Sequence 89165, A
327	23	74.2	438	7	US-11-056-355B-92921	Sequence 92921, A
328	23	74.2	439	7	US-11-056-355B-6673	Sequence 6673, Ap
329	23	74.2	443	7	US-11-056-355B-16814	Sequence 16814, A
330	23	74.2	443	7	US-11-056-355B-6672	Sequence 6672, Ap
331	23	74.2	443	7	US-11-056-355B-16813	Sequence 16813, A
332	23	74.2	445	7	US-11-056-355B-9811	Sequence 9811, Ap
333	23	74.2	451	7	US-11-056-355B-24479	Sequence 24479, A
334	23	74.2	451	7	US-11-056-355B-76335	Sequence 76335, A
335	23	74.2	451	7	US-11-056-355B-80330	Sequence 80330, A
336	23	74.2	451	7	US-11-056-355B-92914	Sequence 92914, A
337	23	74.2	454	7	US-11-056-355B-42310	Sequence 42310, A
338	23	74.2	454	7	US-11-056-355B-89164	Sequence 89164, A
339	23	74.2	454	7	US-11-056-355B-92920	Sequence 92920, A
340	23	74.2	455	6	US-10-449-902-17311	Sequence 17311, A
341	23	74.2	458	7	US-11-056-355B-57434	Sequence 57434, A
342	23	74.2	458	7	US-11-330-403-2460	Sequence 2460, Ap
343	23	74.2	458	7	US-11-330-403-10921	Sequence 10921, A
344	23	74.2	458	7	US-11-330-403-13700	Sequence 13700, A
345	23	74.2	459	6	US-10-449-902-45879	Sequence 45879, A
346	23	74.2	459	6	US-10-449-902-56696	Sequence 56696, A
347	23	74.2	461	6	US-10-449-902-33487	Sequence 33487, A
348	23	74.2	461	6	US-10-449-902-49645	Sequence 49645, A
349	23	74.2	461	6	US-11-056-355B-24478	Sequence 24478, A
350	23	74.2	461	7	US-11-056-355B-89157	Sequence 89157, A
351	23	74.2	461	7	US-11-056-355B-76834	Sequence 76834, A
352	23	74.2	461	7	US-11-056-355B-80329	Sequence 80329, A
353	23	74.2	461	7	US-11-056-355B-89157	Sequence 89157, A
354	23	74.2	461	7	US-11-056-355B-92913	Sequence 92913, A
355	23	74.2	461	7	US-11-056-355B-4945	Sequence 4945, Ap
356	23	74.2	463	7	US-11-056-355B-55446	Sequence 55446, A
357	23	74.2	464	6	US-10-449-902-47092	Sequence 47092, A
358	23	74.2	464	6	US-11-330-403-3890	Sequence 3890, Ap
359	23	74.2	465	7	US-11-330-403-13730	Sequence 13730, A
360	23	74.2	477	7	US-11-056-355B-89156	Sequence 89156, A
361	23	74.2	477	7	US-11-056-355B-92912	Sequence 92912, A
362	23	74.2	479	7	US-11-056-355B-24477	Sequence 24477, A
363	23	74.2	479	7	US-11-056-355B-55445	Sequence 55445, A
364	23	74.2	483	7	US-11-056-355B-76833	Sequence 76833, A
365	23	74.2	483	7	US-11-056-355B-55444	Sequence 55444, A
366	23	74.2	494	6	US-10-449-902-54527	Sequence 54527, A
367	23	74.2	496	6	US-10-449-902-55861	Sequence 55861, A
368	23	74.2	499	7	US-11-293-697-3264	Sequence 3264, Ap
369	23	74.2	505	7	US-11-056-355B-4944	Sequence 4944, Ap
370	23	74.2	509	7	US-11-056-355B-45401	Sequence 45401, A
371	23	74.2	521	6	US-10-449-902-16414	Sequence 16414, A
372	23	74.2	533	6	US-11-056-355B-17703	Sequence 17703, A
373	23	74.2	536	6	US-10-953-349-9750	Sequence 9750, Ap
374	23	74.2	538	7	US-11-330-403-13698	Sequence 13698, A
375	23	74.2	539	7	US-11-330-403-13702	Sequence 13702, A
376	23	74.2	540	6	US-10-449-902-41448	Sequence 41448, A
377	23	74.2	540	6	US-10-449-902-45400	Sequence 45400, A
378	23	74.2	540	6	US-11-056-355B-45400	Sequence 45400, A
379	23	74.2	550	6	US-10-538-066-757	Sequence 757, App
380	23	74.2	550	7	US-11-330-403-13266	Sequence 13266, A
381	23	74.2	563	6	US-10-449-902-43397	Sequence 43397, A
382	23	74.2	565	7	US-11-330-403-16668	Sequence 16668, A
383	23	74.2	575	6	US-10-511-937-2625	Sequence 2625, Ap
384	23	74.2	576	6	US-10-449-902-55695	Sequence 55695, A
385	23	74.2	579	7	US-11-056-355B-17702	Sequence 17702, A
386	23	74.2	593	7	US-11-269-167-2	Sequence 2, Appli
387	23	74.2	597	7	US-11-311-778-16	Sequence 16, Appl
388	23	74.2	620	6	US-10-511-937-2959	Sequence 2959, A
389	23	74.2	626	6	US-10-953-349-7804	Sequence 7804, Ap
390	23	74.2	631	7	US-11-330-403-18087	Sequence 18087, A

391	23	74.2	652	6	US-10-953-349-7803	Sequence 7803, Ap
392	23	74.2	699	6	US-10-196-749-138	Sequence 138, App
393	23	74.2	699	7	US-11-258-349-2	Sequence 2, Appl1
394	23	74.2	711	6	US-10-449-902-45958	Sequence 45958, A
395	23	74.2	718	6	US-10-449-902-44584	Sequence 44584, A
396	23	74.2	718	6	US-11-330-403-12115	Sequence 10215, A
397	23	74.2	783	7	US-11-330-403-4599	Sequence 4599, App
398	23	74.2	815	6	US-10-553-436-183	Sequence 183, App
399	23	74.2	818	6	US-10-449-902-51784	Sequence 51784, A
400	23	74.2	875	6	US-10-471-571A-4802	Sequence 4802, Ap
401	23	74.2	932	6	US-10-563-665-15	Sequence 15, Appl1
402	23	74.2	952	7	US-11-221-332-86	Sequence 86, Appl1
403	23	74.2	1007	7	US-11-273-537-5	Sequence 5, Appl1
404	23	74.2	1217	6	US-10-553-436-160	Sequence 160, App
405	23	74.2	1255	6	US-10-538-066-364	Sequence 364, App
406	23	74.2	1255	7	US-11-223-945-43	Sequence 43, Appl1
407	23	74.2	1255	7	US-11-105-233-130	Sequence 190, App
408	23	74.2	1255	7	US-11-323-049-9	Sequence 9, Appl1
409	23	74.2	1255	7	US-11-323-964-9	Sequence 9, Appl1
410	23	74.2	1255	7	US-11-406-679-6	Sequence 6, Appl1
411	23	74.2	1259	7	US-11-223-945-40	Sequence 40, Appl1
412	23	74.2	1273	7	US-11-056-355B-96451	Sequence 96451, A
413	23	74.2	1308	7	US-11-248-122-2	Sequence 2, Appl1
414	23	74.2	1308	7	US-11-248-936-15	Sequence 15, Appl1
415	23	74.2	1312	7	US-11-056-355B-66450	Sequence 96450, A
416	23	74.2	1323	6	US-10-530-187-289	Sequence 289, App
417	23	74.2	1325	6	US-10-530-187-288	Sequence 288, App
418	23	74.2	1327	7	US-11-221-333-48	Sequence 48, Appl1
419	23	74.2	1330	7	US-11-056-355B-96449	Sequence 96449, A
420	23	74.2	1375	6	US-10-553-436-181	Sequence 181, App
421	23	74.2	1433	7	US-11-056-355B-85575	Sequence 85575, A
422	23	74.2	1439	7	US-11-056-355B-85574	Sequence 85574, A
423	23	74.2	1444	7	US-11-070-573-10	Sequence 10, Appl1
424	23	74.2	1504	7	US-11-056-355B-85573	Sequence 85573, A
425	23	74.2	1533	7	US-11-330-403-9285	Sequence 9285, Ap
426	22	71.0	9	7	US-11-332-378-71	Sequence 71, Appl1
427	22	71.0	9	7	US-11-253-176-133	Sequence 133, App
428	22	71.0	10	6	US-10-542-508-4	Sequence 4, Appl1
429	22	71.0	23	7	US-11-251-734-11	Sequence 11, Appl1
430	22	71.0	48	7	US-11-368-086-73	Sequence 73, Appl1
431	22	71.0	52	7	US-11-115-024-57	Sequence 57, Appl1
432	22	71.0	52	7	US-11-115-024-321	Sequence 321, App
433	22	71.0	58	7	US-11-253-176-10	Sequence 10, Appl1
434	22	71.0	58	7	US-11-253-176-20	Sequence 20, Appl1
435	22	71.0	58	7	US-11-253-176-24	Sequence 24, Appl1
436	22	71.0	62	7	US-11-056-355B-85452	Sequence 85452, A
437	22	71.0	72	6	US-10-953-349-30730	Sequence 30730, A
438	22	71.0	72	7	US-11-056-355B-61292	Sequence 61292, A
439	22	71.0	78	6	US-10-524-434-594	Sequence 594, App
440	22	71.0	79	7	US-11-056-355B-3854	Sequence 3854, Ap
441	22	71.0	86	6	US-10-449-902-36450	Sequence 36450, A
442	22	71.0	86	6	US-10-953-349-4392	Sequence 4392, Ap
443	22	71.0	91	6	US-10-449-902-34501	Sequence 34501, A
444	22	71.0	95	6	US-10-489-071-25	Sequence 25, Appl1
445	22	71.0	103	6	US-10-449-902-36190	Sequence 36190, A
446	22	71.0	105	6	US-10-471-571A-4930	Sequence 4930, Ap
447	22	71.0	108	6	US-10-449-902-46934	Sequence 46934, A
448	22	71.0	111	7	US-11-330-123-60	Sequence 60, Appl1
449	22	71.0	114	7	US-11-293-697-3998	Sequence 3998, Ap
450	22	71.0	114	7	US-11-056-355B-34934	Sequence 34934, A
451	22	71.0	114	7	US-11-056-355B-51981	Sequence 51981, A
452	22	71.0	114	7	US-11-056-355B-88046	Sequence 88046, A
453	22	71.0	116	7	US-11-183-814-7	Sequence 7, Appl1
454	22	71.0	118	6	US-10-471-571A-4412	Sequence 4412, Ap
455	22	71.0	118	7	US-11-140-450-9	Sequence 9, Appl1
456	22	71.0	118	7	US-11-330-403-9062	Sequence 9062, Ap
457	22	71.0	119	6	US-10-542-682-12	Sequence 12, Appl1
458	22	71.0	119	6	US-10-542-682-13	Sequence 13, Appl1
459	22	71.0	119	7	US-11-290-249-7	Sequence 7, Appl1
460	22	71.0	121	7	US-11-254-679-26	Sequence 26, Appl1
461	22	71.0	121	7	US-11-183-814-13	Sequence 13, Appl1
462	22	71.0	124	7	US-11-211-917-50	Sequence 50, Appl1
463	22	71.0	124	7	US-11-211-917-96	Sequence 96, Appl1

464

22

71.0

124

7

US-11-140-450-7

Sequence 7, Appl1

465

22

71.0

124

7

US-11-140-450-8

Sequence 8, Appl1

466

22

71.0

124

7

US-11-330-403-11308

Sequence 11308, A

467

22

71.0

130

6

US-10-471-571A-45618

Sequence 45618, A

468

22

71.0

131

6

US-10-449-902-32612

Sequence 32612, A

469

22

71.0

134

7

US-11-293-687-3462

Sequence 3462, Ap

470

22

71.0

138

7

US-11-056-355B-34933

Sequence 34933, A

471

22

71.0

138

7

US-11-056-355B-88045

Sequence 88045, A

472

22

71.0

138

7

US-11-056-355B-107948

Sequence 107948, A

473

22

71.0

138

7

US-11-056-355B-119187

Sequence 119187, A

474

22

71.0

139

6

US-10-953-349-26277

Sequence 26277, A

475

22

71.0

139

7

US-11-056-355B-51980

Sequence 51980, A

476

22

71.0

139

7

US-11-056-355B-68125

Sequence 68125, A

477

22

71.0

140

6

US-10-449-902-34671

Sequence 34671, A

478

22

71.0

141

6

US-10-449-902-34217

Sequence 34217, A

479

22

71.0

144

7

US-11-056-355B-11833

Sequence 11833, A

480

22

71.0

145

6

US-10-953-349-33596

Sequence 33596, A

481

22

71.0

145

6

US-10-953-349-34570

Sequence 34570, A

482

22

71.0

145

6

US-11-056-355B-6611

Sequence 6617, Ap

483

22

71.0

145

7

US-11-056-355B-15650

Sequence 15650, A

484

22

71.0

147

7

US-11-056-355B-107947

Sequence 107947, A

485

22

71.0

147

7

US-11-056-355B-119186

Sequence 119186, A

486

22

71.0

149

6

US-10-953-349-15605

Sequence 15605, A

487

22

71.0

149

6

US-10-953-349-34569

Sequence 34569, A

488

22

71.0

149

6

US-10-449-902-29048

Sequence 29048, A

489

22

71.0

149

6

US-10-449-902-29990

Sequence 29990, A

490

22

71.0

149

7

US-11-056-355B-6616

Sequence 6616, Ap

491

22

71.0

149

7

US-11-056-355B-57736

Sequence 57736, A

492

22

71.0

153

6

US-10-953-349-26549

Sequence 26549, A

493

22

71.0

164

7

US-11-330-403-12576

Sequence 12576, A

494

22

71.0

165

6

US-10-953-349-33595

Sequence 33595, A

495

22

71.0

165

6

US-10-449-902-36380

Sequence 36380, A

496

22

71.0

165

7

US-11-056-355B-15649

Sequence 15649, A

497

22

71.0

166

7

US-11-056-355B-11832

Sequence 11832, A

498

22

71.0

167

7

US-11-056-355B-52838

Sequence 52838, A

499

22

71.0

168

7

US-11-056-355B-4942

Sequence 4942, Ap

500

22

71.0

169

7

US-11-056-355B-34932

Sequence 34932, A

ALIGNMENTS

RESULT 1

US-11-105-233-53

Sequence 53, Application US/1105233

Publication No. US20060134653A1

GENERAL INFORMATION:

APPLICANT: Thiagalingam et al

TITLE OF INVENTION: Differential Expression of Genes in MSI Tumors

FILE REFERENCE: 1657/2001

CURRENT APPLICATION NUMBER: US/11/105,233

CURRENT FILING DATE: 2005-04-13

NUMBER OF SEQ ID NOS: 202

SOFTWARE: FastSeq For Windows Version 4.0

SEQ ID NO 53

TYPE: PRN

LENGTH: 174

US-11-105-233-53

ORGANISM: Homo sapiens

Query Match

Best Local Similarity 90.3%; Score 28; DB 7; Length 174;

Matches 5; Conservativity 1; Mismatches 0; Indels 0; Gaps 0;

RESULT 2

US-11-056-355B-66811

Sequence 66811, Application US/11056355B

```
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nickolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; TITLE OF INVENTION: Polypeptides Encoded Thereby
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 68811
; LENGTH: 225
; TYPE: prt
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(225)
; OTHER INFORMATION: Ceres Seq. ID no. 12614812
US-11-056-355B-68811
```

```
Query Match          90.3%; Score 28; DB 7; Length 225;
Best Local Similarity 83.3%; Pred. No. 36;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 FSRYAR 6
        ||:||||
Db       144 FSKYAR 149
```

```
RESULT 3
US-11-056-355B-106575
; Sequence 106575, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nickolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; TITLE OF INVENTION: Polypeptides Encoded Thereby
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 106575
; LENGTH: 240
; TYPE: prt
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(240)
; OTHER INFORMATION: Ceres Seq. ID no. 13619980
US-11-056-355B-106575
```

```
Query Match          90.3%; Score 28; DB 7; Length 240;
Best Local Similarity 83.3%; Pred. No. 38;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 FSRYAR 6
        ||:||||
Db       156 FSKYAR 161
```

```
RESULT 4
US-11-056-355B-117814
; Sequence 117814, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nickolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
```

```
; TITLE OF INVENTION: Polypeptides Encoded Thereby
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 117814
; LENGTH: 240
; TYPE: prt
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(240)
; OTHER INFORMATION: Ceres Seq. ID no. 13619980
US-11-056-355B-117814
```

```
Query Match          90.3%; Score 28; DB 7; Length 240;
Best Local Similarity 83.3%; Pred. No. 38;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 FSRYAR 6
        ||:||||
Db       156 FSKYAR 161
```

```
RESULT 5
US-11-056-355B-106574
; Sequence 106574, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nickolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; TITLE OF INVENTION: Polypeptides Encoded Thereby
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 106574
; LENGTH: 243
; TYPE: prt
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(243)
; OTHER INFORMATION: Ceres Seq. ID no. 13619979
US-11-056-355B-106574
```

```
Query Match          90.3%; Score 28; DB 7; Length 243;
Best Local Similarity 83.3%; Pred. No. 39;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 FSRYAR 6
        ||:||||
Db       159 FSKYAR 164
```

```
RESULT 6
US-11-056-355B-117813
; Sequence 117813, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nickolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; TITLE OF INVENTION: Polypeptides Encoded Thereby
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
```

```
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 117813
; LENGTH: 243
; TYPE: prt
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(243)
; OTHER INFORMATION: Ceres Seq. ID no. 13619979
US-11-056-355B-117813

Query Match
Best Local Similarity 90.3%; Score 28; DB 7; Length 243;
Best Local Similarity 83.3%; Pred. No. 39;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSRVAR 6
   ||:||||
Db 159 FSKVAR 164

RESULT 7
US-11-056-355B-106573
; Sequence 106573, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nickolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 106573
; LENGTH: 255
; TYPE: prt
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(255)
; OTHER INFORMATION: Ceres Seq. ID no. 13619978
US-11-056-355B-106573

Query Match
Best Local Similarity 90.3%; Score 28; DB 7; Length 255;
Best Local Similarity 83.3%; Pred. No. 41;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSRVAR 6
   ||:||||
Db 171 FSKVAR 176

RESULT 8
US-11-056-355B-117812
; Sequence 117812, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nickolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 117812
; LENGTH: 255
; TYPE: prt
```

```
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(255)
; OTHER INFORMATION: Ceres Seq. ID no. 13619978
US-11-056-355B-117812

Query Match
Best Local Similarity 90.3%; Score 28; DB 7; Length 255;
Best Local Similarity 83.3%; Pred. No. 41;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSRVAR 6
   ||:||||
Db 171 FSKVAR 176

RESULT 9
US-11-056-355B-68810
; Sequence 68810, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nickolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 68810
; LENGTH: 301
; TYPE: prt
; ORGANISM: Trifolium aestivum
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(301)
; OTHER INFORMATION: Ceres Seq. ID no. 12614811
US-11-056-355B-68810

Query Match
Best Local Similarity 90.3%; Score 28; DB 7; Length 301;
Best Local Similarity 83.3%; Pred. No. 48;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSRVAR 6
   ||:||||
Db 220 FSKVAR 225

RESULT 10
US-11-056-355B-68809
; Sequence 68809, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nickolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 68809
; LENGTH: 348
; TYPE: prt
; ORGANISM: Trifolium aestivum
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(348)
; OTHER INFORMATION: Ceres Seq. ID no. 12614810
```

US-11-056-355B-68809

Query Match 90.3%; Score 28; DB 7; Length 348;
Best Local Similarity 83.3%; Pred. No. 56;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FSRYAR 6
|||
Db 267 FSKYAR 272

RESULT 11

US-10-449-902-37988
; Sequence 37988, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A02051-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 37988
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-37988

Query Match 90.3%; Score 28; DB 6; Length 538;
Best Local Similarity 83.3%; Pred. No. 86;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FSRYAR 6
|||
Db 297 FARYAR 302

RESULT 12

US-10-953-349-9971
; Sequence 9971, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 9971
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-9971

Query Match 87.1%; Score 27; DB 6; Length 129;
Best Local Similarity 83.3%; Pred. No. 35;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FSRYAR 6
|||
Db 31 FSRYGR 36

RESULT 13

US-10-953-349-7495
; Sequence 7495, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 7495
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-7495

Query Match 87.1%; Score 27; DB 6; Length 208;
Best Local Similarity 83.3%; Pred. No. 57;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FSRYAR 6
|||
Db 126 FSRYVR 131

RESULT 14

US-10-953-349-2963
; Sequence 2963, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2963
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-2963

Query Match 87.1%; Score 27; DB 6; Length 209;
Best Local Similarity 83.3%; Pred. No. 57;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FSRYAR 6
|||
Db 128 FSRYVR 133

RESULT 15

US-10-953-349-2962
; Sequence 2962, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2962
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana

US-10-953-349-2962

Query Match 87.1%; Score 27; DB 6; Length 216;
Best Local Similarity 83.3%; Pred. No. 59;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FSRVAR 6
|||
Db 135 FSRYVR 140

RESULT 16

US-10-953-349-7494
; Sequence 7494, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 7494
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-7494

Query Match 87.1%; Score 27; DB 6; Length 222;
Best Local Similarity 83.3%; Pred. No. 61;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FSRVAR 6
|||
Db 140 FSRYVR 145

RESULT 17
US-10-953-349-9130
; Sequence 9130, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 9130
; LENGTH: 290
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-9130

Query Match 87.1%; Score 27; DB 6; Length 290;
Best Local Similarity 83.3%; Pred. No. 79;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FSRVAR 6
|||
Db 31 FSRYVR 36

RESULT 18
US-11-174-307B-202
; Sequence 202, Application US/11174307B
; Publication No. US20060143729A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai

APPLICANT: BROVER, Vyacheslav
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES AND POLYPEPTIDES ENCODED THEREBY
; FILE REFERENCE: 2750-1601PUS2
; CURRENT APPLICATION NUMBER: US/11/174,307B
; CURRENT FILING DATE: 2005-06-30
; PRIOR APPLICATION NUMBER: 60/583,671
; PRIOR FILING DATE: 2004-06-30
; PRIOR APPLICATION NUMBER: 60/583,781
; PRIOR FILING DATE: 2004-06-30
; PRIOR APPLICATION NUMBER: 60/583,651
; PRIOR FILING DATE: 2004-06-30
; NUMBER OF SEQ ID NOS: 5544
; SEQ ID NO 202
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: Pfam Name: RRM_1; Pfam Description: RNA recognition motif.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: Pfam Name: zf-CCHC; Pfam Description: Zinc knuckle
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: GI Number: 11346381; NR Description: splicing factor-like
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: GI Number: 9843663; NR Description: splicing factor RS233
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: GI Number: 42571091; NR Description: splicing factor RS233
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: GI Number: 34895952; NR Description: putative splicing
; OTHER INFORMATION: factor-like protein [Oryza sativa (japonica cultivar-group)]
; OTHER INFORMATION: >gi|1528625|dbj|BAB64646.1| putative splicing factor-like protei
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: GI Number: 62650252; NR Description: PREDICTED: similar to
US-11-174-307B-202

Query Match 87.1%; Score 27; DB 7; Length 312;
Best Local Similarity 83.3%; Pred. No. 85;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FSRVAR 6
|||
Db 59 FSRYVR 64

RESULT 19
US-10-449-902-56777
; Sequence 56777, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30

;; PRIOR APPLICATION NUMBER: JP 2002-383870
;; PRIOR FILING DATE: 2002-12-11
;; NUMBER OF SEQ ID NOS: 56791
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO: 56777
;; LENGTH: 323
;; TYPE: PRT
;; ORGANISM: Oryza sativa
US-10-449-902-56777

Query Match 87.1%; Score 27; DB 6; Length 323;
Best Local Similarity 83.3%; Pred. No. 88;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 FSRYAR 6
 |||||
Db 31 FSRYAR 36

RESULT 20

US-10-953-349-21500
; Sequence 21500, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO: 21500
; LENGTH: 334
; TYPE: PRT
; ORGANISM: Glycine max
US-10-953-349-21500

Query Match 87.1%; Score 27; DB 6; Length 334;
Best Local Similarity 83.3%; Pred. No. 91;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 FSRYAR 6
 |||||
Db 57 FSRYAR 62

RESULT 21

US-10-449-902-34829
; Sequence 34829, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 34829
; LENGTH: 338
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-34829

Query Match 87.1%; Score 27; DB 6; Length 338;

Best Local Similarity 83.3%; Pred. No. 92;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 FSRYAR 6
 |||||
Db 38 FSRYAR 43

RESULT 22

US-10-953-349-21499
; Sequence 21499, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO: 21499
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Glycine max
US-10-953-349-21499

Query Match 87.1%; Score 27; DB 6; Length 353;
Best Local Similarity 83.3%; Pred. No. 96;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 FSRYAR 6
 |||||
Db 76 FSRYAR 81

RESULT 23

US-10-449-902-33374
; Sequence 33374, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 33374
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-33374

Query Match 87.1%; Score 27; DB 6; Length 374;
Best Local Similarity 83.3%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 FSRYAR 6
 |||||
Db 345 FSRYAR 350

RESULT 24

US-10-449-902-45028
; Sequence 45028, Application US/10449902

```
/ Publication No. US20060123505A1
/ GENERAL INFORMATION:
/ APPLICANT: National Institute of Agrobiological Sciences.
/ APPLICANT: Bio-oriented Technology Research Advancement Institution.
/ APPLICANT: The Institute of Physical and Chemical Research.
/ APPLICANT: Foundation for Advancement of International Science.
/ TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
/ FILE REFERENCE: MOA-A0205Y1-US
/ CURRENT APPLICATION NUMBER: US/10/449,902
/ CURRENT FILING DATE: 2003-05-29
/ PRIOR APPLICATION NUMBER: JP 2002-203269
/ PRIOR FILING DATE: 2002-05-30
/ PRIOR APPLICATION NUMBER: JP 2002-383870
/ PRIOR FILING DATE: 2002-12-11
/ NUMBER OF SEQ ID NOS: 56791
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 45028
/ LENGTH: 374
/ TYPE: PRT
/ ORGANISM: Oryza sativa
US-10-449-902-45028

Query Match      87.1%  Score 27; DB 6; Length 374;
Best Local Similarity 83.3%  Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 FSRVAR 6
        |||||
DB      345 FSRVAR 350

RESULT 25
US-10-953-349-21498
/ Sequence 21498, Application US/10953349
/ Publication No. US20060107345A1
/ GENERAL INFORMATION:
/ APPLICANT: ALEXANDROV, Nickolai et al.
/ TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
/ FILE REFERENCE: 2750-1579PUS2
/ CURRENT APPLICATION NUMBER: US/10/953,349
/ CURRENT FILING DATE: 2004-09-30
/ NUMBER OF SEQ ID NOS: 40252
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 21498
/ LENGTH: 464
/ TYPE: PRT
/ ORGANISM: Glycine max
US-10-953-349-21498

Query Match      87.1%  Score 27; DB 6; Length 464;
Best Local Similarity 83.3%  Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 FSRVAR 6
        |||||
DB      187 FSRVAR 192

RESULT 26
US-11-330-403-7004
/ Sequence 7004, Application US/11330403
/ Publication No. US20060159563A1
/ GENERAL INFORMATION:
/ APPLICANT: Abad, Mark S.
/ TITLE OF INVENTION: Genes and Uses for Plant Improvement
/ FILE REFERENCE: 38-21(53629)B
/ CURRENT APPLICATION NUMBER: US/11/330,403
/ CURRENT FILING DATE: 2006-01-12
/ NUMBER OF SEQ ID NOS: 19250
/ SEQ ID NO 7004
/ LENGTH: 478
/ TYPE: PRT
```

```
/ ORGANISM: Caulobacter crescentus CB15
US-11-330-403-7004

Query Match      87.1%  Score 27; DB 7; Length 478;
Best Local Similarity 83.3%  Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 FSRVAR 6
        |||||
DB      423 FSRVAR 428

RESULT 27
US-11-399-675-7
/ Sequence 7, Application US/11399675
/ Publication No. US20060179513A1
/ GENERAL INFORMATION:
/ APPLICANT: Sticklen, Masomeh B
/ APPLICANT: Magbool, Shahina B
/ APPLICANT: Dale, Bruce E
/ TITLE OF INVENTION: TRANSGENIC PLANTS CONTAINING LIGNINASE AND CELLULOSE WHICH
/ FILE REFERENCE: MSU 4.1-814
/ CURRENT APPLICATION NUMBER: US/11/399,675
/ CURRENT FILING DATE: 2006-04-06
/ PRIOR APPLICATION NUMBER: US 60/242,408
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: US 09/981,900
/ PRIOR FILING DATE: 2001-10-18
/ NUMBER OF SEQ ID NOS: 22
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 7
/ LENGTH: 488
/ TYPE: PRT
/ ORGANISM: Actinomyces naeslundii
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (113)..(113)
/ OTHER INFORMATION: The 'Xaa' at location 113 stands for Leu.
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (148)..(148)
/ OTHER INFORMATION: The 'Xaa' at location 148 stands for Asp, Gly, Ala, or Val.
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (316)..(316)
/ OTHER INFORMATION: The 'Xaa' at location 316 stands for Tyr, Cys, Ser, or Phe.
US-11-399-675-7

Query Match      87.1%  Score 27; DB 7; Length 488;
Best Local Similarity 83.3%  Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 FSRVAR 6
        |||||
DB      159 FSRVAR 164

RESULT 28
US-11-353-390A-7
/ Sequence 7, Application US/11353390A
/ Publication No. US20060185036A1
/ GENERAL INFORMATION:
/ APPLICANT: Sticklen, Masomeh B
/ APPLICANT: Magbool, Shahina B
/ APPLICANT: Dale, Bruce E
/ TITLE OF INVENTION: TRANSGENIC PLANTS CONTAINING LIGNINASE AND CELLULOSE WHICH
/ FILE REFERENCE: MSU 4.1-806
/ CURRENT APPLICATION NUMBER: US/11/353,390A
/ CURRENT FILING DATE: 2006-02-14
/ PRIOR APPLICATION NUMBER: US 60/242,408
/ PRIOR FILING DATE: 2000-10-20
```

```

; PRIOR APPLICATION NUMBER: US 09/981,900
; PRIOR FILING DATE: 2001-10-18
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 7
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Actinomyces naeslundii
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (113)..(113)
; OTHER INFORMATION: The 'Xaa' at location 113 stands for Leu.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (148)..(148)
; OTHER INFORMATION: The 'Xaa' at location 148 stands for Asp, Gly, Ala, or Val.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (316)..(316)
; OTHER INFORMATION: The 'Xaa' at location 316 stands for Tyr, Cys, Ser, or Phe.
US-11-353-390A-7
```

```
Query Match      87.1%; Score 27; DB 7; Length 488;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Oy      1 FSRVAR 6
        |||||
Db      159 FERYAR 164
```

```

RESULT 29
US-11-354-310A-7
; Sequence 7, Application US/11354310A
; Publication No. US20060185037A1
; GENERAL INFORMATION:
; APPLICANT: Sticklen, Masomeh B
; APPLICANT: Magbool, Shahina B
; TITLE OF INVENTION: TRANSGENIC PLANTS CONTAINING LIGNINASE AND CELLULOSE WHICH
; FILE REFERENCE: MSU 4.1-805
; CURRENT APPLICATION NUMBER: US/11/354,310A
; CURRENT FILING DATE: 2006-02-14
; PRIOR APPLICATION NUMBER: US 60/242,408
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 09/981,900
; PRIOR FILING DATE: 2001-10-18
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 7
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Actinomyces naeslundii
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (113)..(113)
; OTHER INFORMATION: The 'Xaa' at location 113 stands for Leu.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (148)..(148)
; OTHER INFORMATION: The 'Xaa' at location 148 stands for Asp, Gly, Ala, or Val.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (316)..(316)
; OTHER INFORMATION: The 'Xaa' at location 316 stands for Tyr, Cys, Ser, or Phe.
US-11-354-310A-7
```

```
Query Match      87.1%; Score 27; DB 7; Length 488;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Oy      1 FSRVAR 6
```

```
Db      159 FERYAR 164
```

```

RESULT 30
US-11-174-307B-4272
; Sequence 4272, Application US/11174307B
; Publication No. US20060143729A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai
; APPLICANT: BROVER, Vyacheslav
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES AND POLYPEPTIDES ENCODED THEREBY
; FILE REFERENCE: 2750-1601PUS2
; CURRENT APPLICATION NUMBER: US/11/174,307B
; CURRENT FILING DATE: 2005-06-30
; PRIOR APPLICATION NUMBER: 60/583,671
; PRIOR FILING DATE: 2004-06-30
; PRIOR APPLICATION NUMBER: 60/583,781
; PRIOR FILING DATE: 2004-06-30
; PRIOR APPLICATION NUMBER: 60/583,651
; PRIOR FILING DATE: 2004-06-30
; NUMBER OF SEQ ID NOS: 5544
; SEQ ID NO 4272
; LENGTH: 541
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: Pfam Name: ABC1; Pfam Description: ABC1 family
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: GI Number: 50935257; NR Description: ABC1 family protein
; OTHER INFORMATION: kinase-like protein [Oryza sativa (japonica cultivar-group)]
; OTHER INFORMATION: >gi|2361723|dbj|BAC20904.1| ABC1 family protein kinase-like
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: GI Number: 24215277; NR Description: ABC1 family protein
; OTHER INFORMATION: protein [Leptospira interrogans serovar Lai str. 56601]
; OTHER INFORMATION: >gi|24196371|gb|AA49776.1| ABC1 family protein kinase [Leptospira
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: GI Number: 45657281; NR Description: ubiquinone biosynthesis
; OTHER INFORMATION: protein [Leptospira interrogans serovar Copenhagen str. Pliocruz
; OTHER INFORMATION: LI-110] >gi|45600519|gb|AA570004.1| ubiquinone biosynthesis prote
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: GI Number: 41689606; NR Description: COG0661: Predicted
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: GI Number: 42522036; NR Description: Gene product involved in
; OTHER INFORMATION: ubiquinone production. [Bdellovibrio bacteriovorus HD100]
; OTHER INFORMATION: >gi|39574567|emb|CAE78409.1| Gene product involved in ubiquinone
US-11-174-307B-4272
```

```
Query Match      87.1%; Score 27; DB 7; Length 541;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy      1 FSRVAR 6
        |||||
Db      50 FSRVARQ 55
```

```
RESULT 31
US-10-449-902-33781
; Sequence 33781, Application US/10449902
```

```
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 33781
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Oryza sativa
; US-10-449-902-33781

Query Match      83.9%; Score 26; DB 6; Length 107;
Best Local Similarity 83.3%; Pred. No. 50;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 FSRVAR 6      |||||
DB      88 FSRMAR 93

RESULT 32
US-11-056-355B-581
; Sequence 581, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nikolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; TITLE OF INVENTION: Polypeptides Encoded Thereby
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 581
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(109)
; OTHER INFORMATION: Ceres Seq. ID no. 13635683
; US-11-056-355B-581

Query Match      83.9%; Score 26; DB 7; Length 109;
Best Local Similarity 83.3%; Pred. No. 51;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 FSRVAR 6      |||||
DB      49 FPRVAR 54

RESULT 33
US-11-337-300-451
; Sequence 451, Application US/11337300
; Publication No. US20060121580A1
; GENERAL INFORMATION:
; APPLICANT: Crucell Holland B.V.
; APPLICANT: ter Meulen, Jan H.
; APPLICANT: De Kruijf, Cornelis A.
```

```
; APPLICANT: van den Brink, Edward N.
; APPLICANT: Goudsmut, Jaap
; TITLE OF INVENTION: Binding molecules against SARS-coronavirus and uses thereof
; FILE REFERENCE: 0091 WO 00 ORD
; CURRENT APPLICATION NUMBER: US/11/337,300
; CURRENT FILING DATE: 2006-01-20
; NUMBER OF SEQ ID NOS: 478
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 451
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Variable heavy chain of SC03-057
; US-11-337-300-451

Query Match      83.9%; Score 26; DB 7; Length 129;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 FSRVA 5      |||||
DB      31 FSRVA 35

RESULT 34
US-10-449-902-31883
; Sequence 31883, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 31883
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Oryza sativa
; US-10-449-902-31883

Query Match      83.9%; Score 26; DB 6; Length 217;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 FSRVA 5      |||||
DB      79 FSRVA 83

RESULT 35
US-11-056-355B-16531
; Sequence 16531, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nikolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; TITLE OF INVENTION: Polypeptides Encoded Thereby
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
```

```
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 16531
; LENGTH: 219
; TYPE: prt
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)-(219)
; OTHER INFORMATION: Ceres Seq. ID no. 12352826
US-11-056-355B-16531

Query Match      83.9%; Score 26; DB 7; Length 219;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 FSRYA 5
        |||||
Db      13 FSRYA 17

RESULT 36
US-11-056-355B-16851
; Sequence 16851, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nickolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 16851
; LENGTH: 219
; TYPE: prt
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)-(219)
; OTHER INFORMATION: Ceres Seq. ID no. 12566294
US-11-056-355B-16851

Query Match      83.9%; Score 26; DB 7; Length 219;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 FSRYA 5
        |||||
Db      13 FSRYA 17

RESULT 37
US-11-337-300-286
; Sequence 286, Application US/11337300
; Publication No. US20060121580A1
; GENERAL INFORMATION:
; APPLICANT: Crucell Holland B.V.
; APPLICANT: ter Meulen, Jan H.
; APPLICANT: De Krulff, Cornelis A.
; APPLICANT: van den Brink, Edward N.
; APPLICANT: Goudsmit, Jaap
; TITLE OF INVENTION: Binding molecules against SARS-coronavirus and uses thereof
; FILE REFERENCE: 0091 WO 00 00D
; CURRENT APPLICATION NUMBER: US/11/337,300
; CURRENT FILING DATE: 2006-01-20
; NUMBER OF SEQ ID NOS: 478
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 286
; LENGTH: 256
; TYPE: prt
```

```
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: SC03-057
US-11-337-300-286

Query Match      83.9%; Score 26; DB 7; Length 256;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 FSRYA 5
        |||||
Db      32 FSRYA 36

RESULT 38
US-10-449-902-30299
; Sequence 30299, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30299
; LENGTH: 271
; TYPE: prt
; ORGANISM: Oryza sativa
US-10-449-902-30299

Query Match      83.9%; Score 26; DB 6; Length 271;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 FSRYA 5
        |||||
Db      65 FSRYA 69

RESULT 39
US-10-449-902-31725
; Sequence 31725, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 31725
; LENGTH: 271
; TYPE: prt
; ORGANISM: Oryza sativa
US-10-449-902-31725
```

Query Match 83.9%; Score 26; DB 6; Length 271;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FSRYA 5
|||||
65 FSRYA 69

Db 65 FSRYA 69

RESULT 40
US-10-449-902-32135
; Sequence 32135, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agricultural Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A02051-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; PRIOR FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32135
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-32135

Query Match 83.9%; Score 26; DB 6; Length 271;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FSRYA 5
|||||
65 FSRYA 69

Db 65 FSRYA 69

RESULT 41
US-10-449-902-51604
; Sequence 51604, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agricultural Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A02051-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; PRIOR FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 51604
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-51604

Query Match 83.9%; Score 26; DB 6; Length 271;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FSRYA 5

Db 65 FSRYA 69

RESULT 42
US-11-056-355B-7374
; Sequence 7374, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nikolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; TITLE OF INVENTION: Polypeptides Encoded Thereby
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; PRIOR FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 7374
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(271)
; OTHER INFORMATION: Ceres Seq. ID no. 12412613
US-11-056-355B-7374

Query Match 83.9%; Score 26; DB 7; Length 271;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FSRYA 5
|||||
65 FSRYA 69

Db 65 FSRYA 69

RESULT 43
US-11-056-355B-16849
; Sequence 16849, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nikolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; TITLE OF INVENTION: Polypeptides Encoded Thereby
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; PRIOR FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 16849
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(271)
; OTHER INFORMATION: Ceres Seq. ID no. 12566292
US-11-056-355B-16849

Query Match 83.9%; Score 26; DB 7; Length 271;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FSRYA 5
|||||
65 FSRYA 69

Db 65 FSRYA 69

RESULT 44

```
US-11-056-355B-7372
; Sequence 7372, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nikolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; TITLE OF INVENTION: Polypeptides Encoded Thereby
; FILE REFERENCE: 2/50-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; PRIOR FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 7372
; LENGTH: 311
; TYPE: prt
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(311)
; OTHER INFORMATION: Ceres Seq. ID no. 12412611
US-11-056-355B-7372

Query Match          83.9%; Score 26; DB 7; Length 311;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 FSRVA 5
        |||||
Db      105 FSRVA 109

RESULT 45
US-11-330-403-9151
; Sequence 9151, Application US/11330403
; Publication No. US20060159563A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53629)B
; CURRENT APPLICATION NUMBER: US/11/330,403
; CURRENT FILING DATE: 2006-01-12
; NUMBER OF SEQ ID NOS: 19250
; SEQ ID NO 9151
; LENGTH: 334
; TYPE: prt
; ORGANISM: Staphylococcus aureus subsp. aureus Mu50
US-11-330-403-9151

Query Match          83.9%; Score 26; DB 7; Length 334;
Best Local Similarity 83.3%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      1 FSRVA 6
        |||||
Db      103 FSRVA 108

RESULT 46
US-11-330-403-19189
; Sequence 19189, Application US/11330403
; Publication No. US20060159563A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53629)B
; CURRENT APPLICATION NUMBER: US/11/330,403
; CURRENT FILING DATE: 2006-01-12
; NUMBER OF SEQ ID NOS: 19250
; SEQ ID NO 19189
; LENGTH: 334
; TYPE: prt
```

```
; ORGANISM: Staphylococcus aureus subsp. aureus MW2
US-11-330-403-19189

Query Match          83.9%; Score 26; DB 7; Length 334;
Best Local Similarity 83.3%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      1 FSRVA 6
        |||||
Db      103 FSRVA 108

RESULT 47
US-11-404-745-2
; Sequence 2, Application US/11404745
; Publication No. US20060179505A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Keith D.
; TITLE OF INVENTION: TRANSGENIC MICE CONTAINING CHANNEL
; TITLE OF INVENTION: ACTIVATING PROTEASE 1 (CAP1) GENE DISRUPTIONS
; FILE REFERENCE: R-490
; CURRENT APPLICATION NUMBER: US/11/404,745
; CURRENT FILING DATE: 2006-04-14
; PRIOR APPLICATION NUMBER: US/10/109,616
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 60/280,509
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: US 60/311,055
; PRIOR FILING DATE: 2001-08-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 339
; TYPE: prt
; ORGANISM: Mus musculus
US-11-404-745-2

Query Match          83.9%; Score 26; DB 7; Length 339;
Best Local Similarity 83.3%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      1 FSRVA 6
        |||||
Db      146 FSRVA 151

RESULT 48
US-11-359-554-3
; Sequence 3, Application US/11359554
; Publication No. US20060154293A1
; GENERAL INFORMATION:
; APPLICANT: Xiao, Yonghong
; TITLE OF INVENTION: Regulation of Human Proctasin-Like
; TITLE OF INVENTION: Serine Protease
; FILE REFERENCE: 004974.00929
; CURRENT APPLICATION NUMBER: US/11/359,554
; CURRENT FILING DATE: 2006-02-23
; PRIOR APPLICATION NUMBER: US/10/311,591
; PRIOR FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: US 60/213,474
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/277,612
; PRIOR FILING DATE: 2001-03-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 343
; TYPE: prt
; ORGANISM: Homo sapiens
US-11-359-554-3

Query Match          83.9%; Score 26; DB 7; Length 343;
Best Local Similarity 83.3%; Pred. No. 1.6e+02;
```

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FSRVAR 6
 |||||
 Db 146 FSRVAR 151

Search completed: August 29, 2006, 11:29:12
 Job time : 4.75248 secs

RESULT 49
 US-10-471-571A-4440
 ; Sequence 4440, Application US/10471571A
 ; Publication No. US20060115490A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CHIRON SPA
 ; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
 ; FILE REFERENCE: P026927W0
 ; CURRENT APPLICATION NUMBER: US/10/471,571A
 ; CURRENT FILING DATE: 2003-09-12
 ; PRIOR APPLICATION NUMBER: GB-0107661.1
 ; PRIOR FILING DATE: 2001-03-27
 ; NUMBER OF SEQ ID NOS: 5642
 ; SOFTWARE: Seqwin99, version 1.03
 ; SEQ ID NO 4440
 ; LENGTH: 348
 ; TYPE: PRT
 ; ORGANISM: Staphylococcus aureus
 ; FEATURE:
 ; NAME/KEY: MISC FEATURE
 ; LOCATION: (1)..(348)
 ; OTHER INFORMATION: conserved hypothetical
 US-10-471-571A-4440

Query Match 83.9%; Score 26; DB 6; Length 348;
 Best Local Similarity 83.3%; Pred. No. 1.6e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FSRVAR 6
 |||||
 Db 117 FSRVAR 122

RESULT 50
 US-11-333-747A-22
 ; Sequence 22, Application US/11333747A
 ; Publication No. US20060160159A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Handfield, Martin
 ; APPLICANT: Hillman, Jeffrey
 ; APPLICANT: Proguiske-Fox, Ann
 ; TITLE OF INVENTION: Identification of Actinobacillus actinomycetemcomitans
 ; TITLE OF INVENTION: Antigens for use in the diagnosis, treatment, and monitoring
 ; TITLE OF INVENTION: of Periodontal Diseases
 ; FILE REFERENCE: MBHB01-662C
 ; CURRENT APPLICATION NUMBER: US/11/333,747A
 ; CURRENT FILING DATE: 2006-01-17
 ; PRIOR APPLICATION NUMBER: US 09/995493
 ; PRIOR FILING DATE: 2001-11-28
 ; PRIOR APPLICATION NUMBER: PCT/US02/37235
 ; PRIOR FILING DATE: 2002-11-20
 ; NUMBER OF SEQ ID NOS: 234
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 22
 ; LENGTH: 365
 ; TYPE: PRT
 ; ORGANISM: Actinobacillus actinomycetemcomitans
 US-11-333-747A-22

Query Match 83.9%; Score 26; DB 7; Length 365;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSRVA 5
 |||||
 Db 278 FSRVA 282

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2006, 05:46:47 ; Search time 45.7426 Seconds
(without alignments)
69.968 Million cell updates/sec

Title: US-10-541-343-3

Perfect score: 38

Sequence: 1 WMLSAFS 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

2589679

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database :
1: A_Geneseq.8:*
2: geneseqp1980s:*
3: geneseqp1990s:*
4: geneseqp2000s:*
5: geneseqp2001s:*
6: geneseqp2002s:*
7: geneseqp2003as:*
8: geneseqp2003bs:*
9: geneseqp2004s:*
10: geneseqp2005s:*
11: geneseqp2006s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	100.0	7	ADQ90197	Adq90197 WMLSAFS d
2	38	100.0	13	ADQ90200	Adq90200 MNTF 13 m
3	38	100.0	33	AAW59046	Aaw59046 Human MNT
4	38	100.0	33	AAO29914	Aao29914 Human mot
5	38	100.0	33	ADQ90195	Adq90195 MNTF1-F6
6	35	92.1	1760	ABBS3367	Abbs3367 Drosophila
7	35	92.1	1760	ABG70020	Abg70020 Larval vl
8	33	86.8	17	ADX56419	Adx56419 Cardiov
9	33	86.8	35	AAAB29781	Aab29781 Human MSF
10	33	86.8	61	AAAB4007	Aab4007 Human can
11	33	86.8	538	AAO18834	Aao18834 3' cartil
12	33	86.8	902	AAAB29778	Aab29778 Human MSF
13	33	86.8	933	ADQ67713	Adq67713 Novel hum
14	33	86.8	981	ADZ03664	Adz03664 PRG4-Lub:
15	33	86.8	1007	ADZ03668	Adz03668 PRG4-Lub:
16	33	86.8	1038	ADZ03672	Adz03672 PRG4-Lub:
17	33	86.8	1069	ADZ03676	Adz03676 PRG4-Lub:
18	33	86.8	1100	ADZ03680	Adz03680 PRG4-Lub:
19	33	86.8	1270	ADK67911	Adk67911 Human ext
20	33	86.8	1299	AAK24322	Aam24322 Human EST
21	33	86.8	1311	ADK67912	Adk67912 Human ext
22	33	86.8	1330	ADK65819	Adk65819 Angiogene
23	33	86.8	1404	AAK26049	Aar26049 MSF precu

24	33	86.8	1404	AAAB29773	Aab29773 Human meg
25	33	86.8	1404	AAAB60568	Aab60568 Human meg
26	33	86.8	1404	ADK65839	Adk65839 Angiogene
27	33	86.8	1404	ADM98014	Adm98014 Human meg
28	33	86.8	1404	ADK69335	Adk69335 Human meg
29	33	86.8	1404	ADY15658	Ady15658 Human hep
30	33	86.8	1404	ADK6477	Adk6477 Human lib
31	33	86.8	1404	AEF54256	Aef54256 Megakaryo
32	33	86.8	1404	AEF89839	Aef89839 Human meg
33	33	86.8	1415	AAU32262	Aau32262 Novel hum
34	32	84.2	31	AAAB2730	Aab2730 Human imm
35	32	84.2	81	ABP04103	Abp04103 Human ORF
36	32	84.2	103	ADC95262	Adc95262 E. faeciu
37	32	84.2	211	AAU25590	Aau25590 Human G p
38	32	84.2	243	ADK09699	Adk09699 Human pro
39	32	84.2	262	ABO73613	Ab073613 Pseudomon
40	32	84.2	280	AAAG81234	Aag81234 Mycobacte
41	32	84.2	289	AAAM51647	Aam51647 Human GPC
42	32	84.2	289	ADK029555	Adk029555 Mouse GPC
43	32	84.2	289	ADK029554	Adk029554 Human GPC
44	32	84.2	323	AAU25612	Aau25612 Human G p
45	32	84.2	323	ABBO6257	Abbo6257 Human G p
46	32	84.2	371	ABU10905	Abu10905 Human sec
47	32	84.2	437	ABO60886	Ab060886 Klebsiell
48	32	84.2	561	ADF70484	Adf70484 Orphan re
49	32	84.2	1180	ADU02401	Adu02401 Novel hum
50	32	84.2	1923	ABG21342	Abg21342 Novel hum
51	31	81.6	13	ABK23728	Abk23728 Phaneiroch
52	31	81.6	76	AAAB94607	Aaab94607 Human rep
53	31	81.6	76	ADK09815	Adk09815 Altiolococ
54	31	81.6	91	ABP52882	Abp52882 Human imm
55	31	81.6	93	AAAB5209	Aaab5209 Human imm
56	31	81.6	146	AAAB54191	Aab54191 Human pan
57	31	81.6	193	AGK30491	Agk30491 Arabidops
58	31	81.6	199	ADU17777	Adu17777 Reverse t
59	31	81.6	204	AAAG30490	Aag30490 Arabidops
60	31	81.6	229	AAAG6722	Aag6722 Arabidops
61	31	81.6	229	AAAG45332	Aag45332 Arabidops
62	31	81.6	263	AAU082738	Aau082738 Amino aci
63	31	81.6	265	AAAG30489	Aag30489 Arabidops
64	31	81.6	290	AAAG6721	Aag6721 Arabidops
65	31	81.6	290	AAAG45331	Aag45331 Arabidops
66	31	81.6	299	AAAG45330	Aag45330 Arabidops
67	31	81.6	299	AAAG6720	Aag6720 Arabidops
68	31	81.6	299	ADN74439	Adn74439 Thale cre
69	31	81.6	305	ADK74105	Adk74105 Plant ful
70	31	81.6	362	AAAB81100	Aab81100 Sequence
71	31	81.6	362	ABK70453	Abk70453 Drosophila
72	31	81.6	371	AAU79553	Aau79553 P. chryso
73	31	81.6	397	ADK213153	Adk213153 Human can
74	30	78.9	50	ABO55588	Ab055588 Human gen
75	30	78.9	117	ADK40178	Adk40178 HIV Rev p
76	30	78.9	119	ADK94073	Adk94073 E. faeciu
77	30	78.9	152	ABO14240	Ab014240 Novel hum
78	30	78.9	152	ADK68647	Adk68647 Human sec
79	30	78.9	152	ADK60938	Adk60938 Human sec
80	30	78.9	155	ADK09737	Adk09737 Altiolococ
81	30	78.9	158	AAAG72847	Aag72847 Human oif
82	30	78.9	174	ADK57140	Adk57140 Human sec
83	30	78.9	174	ADK41004	Adk41004 Human sec
84	30	78.9	174	ABO14025	Ab014025 Novel hum
85	30	78.9	174	ADK37949	Adk37949 Human sec
86	30	78.9	174	ADK78431	Adk78431 Human sec
87	30	78.9	174	ADK60722	Adk60722 Human sec
88	30	78.9	175	AAAG72846	Aag72846 Human oif
89	30	78.9	176	AAAG72845	Aag72845 Human oif
90	30	78.9	182	ADK23223	Adk23223 Bacteri
91	30	78.9	191	ADK95755	Adk95755 E. faeciu
92	30	78.9	230	ABO81807	Ab081807 Pseudomon
93	30	78.9	244	AEBA17722	Aeb17722 L. pneumo
94	30	78.9	246	AEBS8464	Aeb38464 L. pneumo
95	30	78.9	273	ADN20738	Adn20738 Bacteri
96	30	78.9	320	ADK93502	Adk93502 Pavlova 1

97	30	78.9	320	9	ADM11906	Adw11906	Pavlova 1	170	29	76.3	335	6	ABU82132	Abu82132	Novel hum
98	30	78.9	320	9	AED95183	Aed95183	Pavlova 1	171	29	76.3	335	6	ABU72312	Abu72312	Human PRO
99	30	78.9	321	6	ADA36340	Ada36340	Acinetoba	172	29	76.3	335	6	ABO33618	AbO33618	Novel hum
100	30	78.9	332	4	ABAB8503	Abab8503	Haemophil	173	29	76.3	335	6	ABJ72440	AbJ72440	Human PRO
101	30	78.9	332	5	AAU91434	Aau91434	Haemophil	174	29	76.3	335	6	ABO34335	AbO34335	Human sec
102	30	78.9	359	7	ABO67049	AbO67049	Klebsiell	175	29	76.3	335	7	ABO44471	AbO44471	Human sec
103	30	78.9	420	6	ABP56696	AbP56696	Human int	176	29	76.3	335	7	ABO33495	AbO33495	Novel hum
104	30	78.9	464	8	ADH39760	Adh39760	Streptomy	177	29	76.3	335	7	ABU72142	AbJ72142	Human mem
105	30	78.9	481	7	ABM90405	Abm90405	Rice abio	178	29	76.3	335	7	ADH83660	Adh83660	Novel hum
106	30	78.9	486	7	ABM66642	Abm66642	Rice abio	179	29	76.3	335	7	ADH80766	Adh80766	Novel hum
107	30	78.9	526	4	AA934482	Aa934482	Human pro	180	29	76.3	335	7	ADH73307	Adh73307	Novel hum
108	30	78.9	526	5	ABB97533	Abb97533	Novel hum	181	29	76.3	335	7	ADH78389	Adh78389	Novel hum
109	30	78.9	628	7	ABO75995	AbO75995	Pseudomon	182	29	76.3	335	7	ADH85037	Adh85037	Human PRO
110	30	78.9	736	8	ADX68450	Adx68450	Plant ful	183	29	76.3	335	7	ADH78143	Adh78143	Novel hum
111	30	78.9	818	8	ADX71975	Adx71975	Plant ful	184	29	76.3	335	7	ADH87209	Adh87209	Human PRO
112	30	78.9	907	4	ABBB62840	Abb62840	Drosophil	185	29	76.3	335	7	ADH84791	Adh84791	Human PRO
113	30	78.9	907	4	ABBA4576	Abba4576	Sleep dis	186	29	76.3	335	7	ADH83906	Adh83906	Novel hum
114	30	78.9	1321	8	ADN23423	Adn23423	Bacteri-al	187	29	76.3	335	7	ADH73061	Adh73061	Novel hum
115	30	78.9	1372	8	ADN23424	Adn23424	Bacteri-al	188	29	76.3	335	7	ADH71999	Adh71999	Human PRO
116	30	78.9	2250	9	ADY27516	Ady27516	HCCV-NL63	189	29	76.3	335	7	ADH7209	Adh7209	Human PRO
117	30	78.9	4055	9	AEA06399	Aea06399	EMCR coro	190	29	76.3	335	7	ADH7209	Adh7209	Human PRO
118	30	78.9	4060	9	ADY27513	Ady27513	Amrno aci	191	29	76.3	335	7	ADH7209	Adh7209	Human PRO
119	30	78.9	6738	9	ADY27514	Ady27514	Amrno aci	192	29	76.3	335	7	ADH7209	Adh7209	Human PRO
120	30	78.9	6738	9	ADY27515	Ady27515	Amrno aci	193	29	76.3	335	7	ADH7209	Adh7209	Human PRO
121	29	76.3	58	5	ABP11194	Abp11194	Human ORF	194	29	76.3	335	7	ADH7209	Adh7209	Human PRO
122	29	76.3	60	3	AAAB38320	AaB38320	Human sec	195	29	76.3	335	7	ADH7209	Adh7209	Human PRO
123	29	76.3	87	4	AAU17729	Aau17729	Novel hum	196	29	76.3	335	7	ADH7209	Adh7209	Human PRO
124	29	76.3	87	7	ADG41109	Adg41109	Human res	197	29	76.3	335	7	ADH7209	Adh7209	Human PRO
125	29	76.3	87	7	ADP96883	Adp96883	Human res	198	29	76.3	335	7	ADH7209	Adh7209	Human PRO
126	29	76.3	106	5	ABAB89390	Abb89390	Human pol	199	29	76.3	335	7	ADH7209	Adh7209	Human PRO
127	29	76.3	113	8	ADY06261	Ady06261	Plant ful	200	29	76.3	335	7	ADH7209	Adh7209	Human PRO
128	29	76.3	113	8	ADY06261	Ady06261	Plant ful	201	29	76.3	335	7	ADH7209	Adh7209	Human PRO
129	29	76.3	114	8	ADY06261	Ady06261	Plant ful	202	29	76.3	335	7	ADH7209	Adh7209	Human PRO
130	29	76.3	117	8	ADY05513	Ady05513	Plant ful	203	29	76.3	335	7	ADH7209	Adh7209	Human PRO
131	29	76.3	127	2	AAH88723	Aah88723	Secreted	204	29	76.3	335	7	ADH7209	Adh7209	Human PRO
132	29	76.3	127	4	ABBS0490	Abb50490	Human sec	205	29	76.3	335	7	ADH7209	Adh7209	Human PRO
133	29	76.3	127	6	ABO44747	AbO44747	Novel hum	206	29	76.3	335	7	ADH7209	Adh7209	Human PRO
134	29	76.3	127	6	ABO26227	AbO26227	Human pro	207	29	76.3	335	7	ADH7209	Adh7209	Human PRO
135	29	76.3	129	5	ABP06040	Abp06040	Human ORF	208	29	76.3	335	7	ADH7209	Adh7209	Human PRO
136	29	76.3	132	4	AAO04045	Aao04045	Human pol	209	29	76.3	335	7	ADH7209	Adh7209	Human PRO
137	29	76.3	146	6	ABU18616	Abu18616	Protein e	210	29	76.3	335	7	ADH7209	Adh7209	Human PRO
138	29	76.3	146	7	ADFS9661	Adf59661	Human pol	211	29	76.3	335	7	ADH7209	Adh7209	Human PRO
139	29	76.3	186	8	ADX90922	Adx90922	Plant ful	212	29	76.3	335	7	ADH7209	Adh7209	Human PRO
140	29	76.3	200	4	AAO06824	Aao06824	Human pol	213	29	76.3	335	7	ADH7209	Adh7209	Human PRO
141	29	76.3	218	4	AAAB95287	Aab95287	Human pro	214	29	76.3	335	7	ADH7209	Adh7209	Human PRO
142	29	76.3	222	9	ADY65804	Ady65804	S. mansoni	215	29	76.3	335	7	ADH7209	Adh7209	Human PRO
143	29	76.3	238	7	ADFS8723	Adf58723	Human pol	216	29	76.3	335	7	ADH7209	Adh7209	Human PRO
144	29	76.3	243	2	AAW55241	Aaw55241	H. pylori	217	29	76.3	335	7	ADH7209	Adh7209	Human PRO
145	29	76.3	250	9	ADM26637	Adm26637	Fructo-col	218	29	76.3	335	7	ADH7209	Adh7209	Human PRO
146	29	76.3	252	8	ADY06260	Ady06260	Plant ful	219	29	76.3	335	7	ADH7209	Adh7209	Human PRO
147	29	76.3	272	2	AAW71530	Aaw71530	Helicobac	220	29	76.3	335	7	ADH7209	Adh7209	Human PRO
148	29	76.3	279	2	AAW55471	Aaw55471	H. pylori	221	29	76.3	335	7	ADH7209	Adh7209	Human PRO
149	29	76.3	293	3	AAW54475	Aaw54475	Arabi-dops	222	29	76.3	335	7	ADH7209	Adh7209	Human PRO
150	29	76.3	293	4	ABG06787	Abg06787	Novel hum	223	29	76.3	335	8	ADH7209	Adh7209	Human PRO
151	29	76.3	295	3	AAAG49477	Aag49477	Arabi-dops	224	29	76.3	335	8	ADH7209	Adh7209	Human PRO
152	29	76.3	303	2	AAW37367	Aaw37367	PsdL (Wbp	225	29	76.3	335	8	ADH7209	Adh7209	Human PRO
153	29	76.3	304	3	AAAG44755	Aag44755	Arabi-dops	226	29	76.3	335	8	ADH7209	Adh7209	Human PRO
154	29	76.3	306	3	AAAG49476	Aag49476	Arabi-dops	227	29	76.3	335	8	ADH7209	Adh7209	Human PRO
155	29	76.3	307	3	AAAG32966	Aag32966	Arabi-dops	228	29	76.3	335	8	ADH7209	Adh7209	Human PRO
156	29	76.3	307	3	ABU52708	Abu52708	Human cel	229	29	76.3	335	8	ADH7209	Adh7209	Human PRO
157	29	76.3	335	4	AAIY94899	Aaiy94899	Human sec	230	29	76.3	335	8	ADH7209	Adh7209	Human PRO
158	29	76.3	335	3	AAAB42940	Aab42940	Human ORF	231	29	76.3	335	8	ADH7209	Adh7209	Human PRO
159	29	76.3	335	3	AAAB24423	Aab24423	Human ORF	232	29	76.3	335	8	ADH7209	Adh7209	Human PRO
160	29	76.3	335	3	AAIY93376	Aaiy93376	Human pro	233	29	76.3	335	8	ADH7209	Adh7209	Human PRO
161	29	76.3	335	4	AAAB6125	Aab66125	Protein o	234	29	76.3	335	8	ADH7209	Adh7209	Human PRO
162	29	76.3	335	4	AAAB88486	Aab88486	Human mem	235	29	76.3	335	8	ADH7209	Adh7209	Human PRO
163	29	76.3	335	4	ABU52707	Abu52707	Human cel	236	29	76.3	335	8	ADH7209	Adh7209	Human PRO
164	29	76.3	335	5	AAU83676	Aau83676	Human PRO	237	29	76.3	335	8	ADH7209	Adh7209	Human PRO
165	29	76.3	335	5	ABBB84911	Abb84911	Human PRO	238	29	76.3	335	8	ADH7209	Adh7209	Human PRO
166	29	76.3	335	5	ABBB84911	Abb84911	Human PRO	239	29	76.3	335	8	ADH7209	Adh7209	Human PRO
167	29	76.3	335	5	ABBB84911	Abb84911	Human PRO	240	29	76.3	335	8	ADH7209	Adh7209	Human PRO
168	29	76.3	335	5	ABBB84911	Abb84911	Human PRO	241	29	76.3	335	8	ADH7209	Adh7209	Human PRO
169	29	76.3	335	6	ABO33789	AbO33789	Novel hum	242	29	76.3	335	8	ADH7209	Adh7209	Human PRO

243	29	76.3	335	8	ADb85601	Adb85601	Novel	hum	316	29	76.3	624	7	ADb60520	Adbe0520	Rat	Prote
244	29	76.3	335	8	ADb05150	Adbe05150	Human	PRO	317	29	76.3	624	7	ADb60516	Adbe0516	Rat	Prote
245	29	76.3	335	8	ADb75363	Adb75363	Human	PRO	318	29	76.3	624	8	ADb09403	Adbe09403	Human	pro
246	29	76.3	335	8	ADb76907	Adb76907	Novel	hum	319	29	76.3	689	4	ABG00041	Abg00041	Novel	hum
247	29	76.3	335	8	ADb86675	Adb86675	Novel	hum	320	29	76.3	722	10	AEF13718	Aef13718	Aphanotne	
248	29	76.3	335	8	ADb78143	Adb78143	Novel	hum	321	29	76.3	722	9	AEA21035	Aea21035	Novel	hum
249	29	76.3	335	8	ADb77651	Adb77651	Novel	hum	322	29	76.3	806	5	ABP62885	Abp62885	Human	pro
250	29	76.3	335	8	ADb77897	Adb77897	Novel	hum	323	29	76.3	806	7	ADb69027	Adb69027	Human	hea
251	29	76.3	335	8	ADb85355	Adb85355	Novel	hum	324	29	76.3	806	8	ADb65804	Adb65804	Novel	hum
252	29	76.3	335	8	ADb73887	Adb73887	Human	PRO	325	29	76.3	810	8	ADbQ18928	AdbQ18928	Human	sof
253	29	76.3	335	8	ADb74625	Adb74625	Human	PRO	326	29	76.3	845	3	AAQ30758	AAQ30758	Arabi	dops
254	29	76.3	335	8	ADb77153	Adb77153	Novel	hum	327	29	76.3	851	6	ABU37008	Abu37008	Protein	e
255	29	76.3	335	8	ADb85847	Adb85847	Novel	hum	328	29	76.3	859	3	AAQ30756	AAQ30756	Arabi	dops
256	29	76.3	335	8	ADb805396	Adb805396	Human	PRO	329	29	76.3	895	3	AAQ45340	AAQ45340	Arabi	dops
257	29	76.3	335	8	ADb74871	Adb74871	Human	PRO	330	29	76.3	903	3	AAQ45339	AAQ45339	Arabi	dops
258	29	76.3	335	8	ADb86503	Adb86503	Human	sec	331	29	76.3	982	3	AAQ45339	AAQ45339	Arabi	dops
259	29	76.3	335	8	ADbF25814	Adbf25814	Human	sec	332	29	76.3	1065	6	ABU37008	Abu37008	Protein	e
260	29	76.3	335	8	ADbF24713	Adbf24713	Human	sec	333	29	76.3	1818	5	ABP73779	Abp73779	Candida	a
261	29	76.3	335	8	ADbF29449	Adbf29449	Human	sec	334	29	76.3	1878	4	ABB61852	Abb61852	Drosophi	l
262	29	76.3	335	8	ADb86980	Adb86980	Human	sec	335	29	76.3	2005	4	ABG24464	ABG24464	Novel	hum
263	29	76.3	335	8	ADb805683	Adb805683	Novel	hum	336	29	76.3	4101	8	ADN23857	Adn23857	Bacterial	
264	29	76.3	335	8	ADbG27237	Adbg27237	Human	PRO	337	29	76.3	4101	8	ADN23856	Adn23856	Bacterial	
265	29	76.3	335	8	ADbG11300	Adbg11300	Novel	hum	338	29	76.3	4924	4	AAb70968	Aab70968	S. spinos	
266	29	76.3	335	8	ADbG12079	Adbg12079	Novel	hum	339	29	76.3	4928	2	AAy39300	Aay39300	Spond	a po
267	29	76.3	335	8	ADbP94636	Adbp94636	Novel	hum	340	29	76.3	4933	6	ABP57681	Abp57681	Saccharop	
268	29	76.3	335	8	ADbG06732	Adbg06732	Human	PRO	341	28	73.7	25	8	ABO54055	AbO54055	Human	gen
269	29	76.3	335	8	ADbH03018	Adbh03018	Human	sec	342	28	73.7	52	8	ADb87749	Adb87749	Plant	ful
270	29	76.3	335	8	ADbH03972	Adbh03972	Human	sec	343	28	73.7	60	4	AAm37179	Aam37179	Peptide	#
271	29	76.3	335	8	ADbH03495	Adbh03495	Human	sec	344	28	73.7	61	4	AAQ77571	Aaq77571	Human	col
272	29	76.3	335	8	ADbH39076	Adbh39076	Novel	hum	345	28	73.7	63	4	ABG08861	Abg08861	Novel	hum
273	29	76.3	335	8	ADbH43623	Adbh43623	Human	PRO	346	28	73.7	71	3	AAQ35290	Aaq35290	Zea	mays
274	29	76.3	335	8	ADbG34166	Adbg34166	Novel	hum	347	28	73.7	71	3	AAQ35299	Aaq35299	Zea	mays
275	29	76.3	335	8	ADbH04449	Adbh04449	Human	sec	348	28	73.7	73	3	AAQ18848	Aaq18848	Zea	mays
276	29	76.3	335	8	ADbI33636	Adbi33636	Human	PRO	349	28	73.7	74	3	ABE29829	AbE29829	Human	sec
277	29	76.3	335	8	ADbH61450	Adbh61450	Human	sec	350	28	73.7	75	2	AAW20820	Aaw20820	H. pylori	
278	29	76.3	335	8	ADbH61450	Adbh61450	Human	sec	351	28	73.7	75	4	ABB41555	Abb41555	Peptide	#
279	29	76.3	335	8	ADbI29891	Adbi29891	Novel	hum	352	28	73.7	75	4	AAm35346	Aam35346	Peptide	#
280	29	76.3	335	8	ADbM27288	Adbm27288	Novel	hum	353	28	73.7	75	4	AAm75234	Aam75234	Human	bon
281	29	76.3	335	8	ADbK82968	Adbk82968	Human	PRO	354	28	73.7	75	4	AAm62426	Aam62426	Human	bra
282	29	76.3	335	8	ADbK66646	Adbk66646	Human	PRO	355	28	73.7	75	4	ABG56996	Abg56996	Human	liv
283	29	76.3	335	8	ADbI91857	Adbi91857	Human	PRO	356	28	73.7	75	5	ABG44887	Abg44887	Human	pep
284	29	76.3	335	8	ADbI94649	Adbi94649	Human	sec	357	28	73.7	76	4	AAU45193	Aau45193	Propionib	
285	29	76.3	335	8	ADbT94363	Adbt94363	Human	PRO	358	28	73.7	76	4	AAU63295	Aau63295	Propionib	
286	29	76.3	335	9	ADbY17816	Adby17816	PRO	polyp	359	28	73.7	76	6	ABM41712	Abm41712	Propionib	
287	29	76.3	335	9	ADbY63337	Adby63337	Human	c10	360	28	73.7	76	6	ABM59814	Abm59814	Propionib	
288	29	76.3	335	9	ADbY79911	Adby79911	Amtno	ac1	361	28	73.7	78	4	AAW24468	Aaw24468	Human	EST
289	29	76.3	335	9	AEbD45036	Aebd45036	Human	sec	362	28	73.7	80	2	AAy60165	Aay60165	Human	EST
290	29	76.3	336	6	ABM67175	Abm67175	Phototrab		363	28	73.7	80	6	ABJ18385	Abj18385	Breast	sp
291	29	76.3	347	3	AAQ44754	Aaq44754	Arabi	dops	364	28	73.7	82	3	AAQ33806	Aaq33806	Arabi	dops
292	29	76.3	349	3	AAQ49475	Aaq49475	Arabi	dops	365	28	73.7	89	5	ABB98392	Abb98392	Cytochrom	
293	29	76.3	350	3	AAQ32965	Aaq32965	Arabi	dops	366	28	73.7	89	5	ABJ10235	Abj10235	Human	lun
294	29	76.3	355	8	ADbT59963	Adbt59963	Plant	pol	367	28	73.7	93	9	AEbD3040	Aebd3040	Hyperimmu	
295	29	76.3	356	5	ABB47847	Abb47847	Listeria		368	28	73.7	98	7	ADbI5994	Adbi5994	G-coupled	
296	29	76.3	357	8	ADbT59952	Adbt59952	Plant	pol	369	28	73.7	98	8	ADbI59933	Adbi59933	Human	G-c
297	29	76.3	359	7	ABO79984	AbO79984	Pseudomon		370	28	73.7	100	8	ABO54783	AbO54783	Human	gen
298	29	76.3	360	4	AAm00105	Aam00105	Tritosepho		371	28	73.7	103	5	ABBI10024	Abbi10024	Human	pro
299	29	76.3	376	8	ADbQ65381	Adbq65381	Novel	hum	372	28	73.7	103	5	ABBI10006	Abbi10006	Human	pro
300	29	76.3	381	3	AAQ32964	Aaq32964	Arabi	dops	373	28	73.7	103	5	ABG64646	Abg64646	Human	alb
301	29	76.3	407	8	ADbY23095	Adby23095	Plant	ful	374	28	73.7	103	5	ABG64647	Abg64647	Human	alb
302	29	76.3	414	8	ADbY09740	Adby09740	Plant	ful	375	28	73.7	103	6	ADbA55367	Adba55367	Human	pro
303	29	76.3	428	8	ADbY09849	Adby09849	Plant	ful	376	28	73.7	103	6	ADbA57056	Adba57056	Human	pro
304	29	76.3	437	5	ADbF91891	Adbf91891	Agarase	I	377	28	73.7	103	6	ADbA57460	Adba57460	Human	sec
305	29	76.3	445	8	ADbI92225	Adbi92225	Urokinase		378	28	73.7	103	6	ADbA41339	Adba41339	Human	sec
306	29	76.3	463	8	ADbJ35098	Adbj35098	Xyltanase		379	28	73.7	103	6	ADbA40909	Adba40909	Human	sec
307	29	76.3	483	5	ABP29825	Abp29825	Streptococ		380	28	73.7	103	6	ABR47825	AbR47825	Human	sec
308	29	76.3	483	8	ADb88149	Adb88149	Streptococ		381	28	73.7	103	6	ABR48062	AbR48062	Human	sec
309	29	76.3	483	8	ADbV74402	Adbv74402	Streptococ		382	28	73.7	103	7	ADbC74494	AdbC74494	Human	sec
310	29	76.3	488	9	ADbV81576	Adbv81576	Streptococ		383	28	73.7	103	7	ADbC74214	AdbC74214	Human	sec
311	29	76.3	488	9	ADbC57468	Adbc57468	L. acidipoc		384	28	73.7	103	7	ADbD38047	Adbd38047	Human	sec
312	29	76.3	495	9	ADbC57264	Adbc57264	L. acidipoc		385	28	73.7	103	7	ADbD37922	Adbd37922	Human	sec
313	29	76.3	545	8	ADbN21402	Adbn21402	Bacterial		386	28	73.7	103	8	ADbL77913	Adbl77913	Albumin	f
314	29	76.3	585	8	ADbN27112	Adbn27112	Bacterial		387	28	73.7	103	8	ADbL77914	Adbl77914	Albumin	f
315	29	76.3	598	8	ADbS29798	Adbs29798	Bacterial		388	28	73.7	107	3	AAb40942	Abb40942	Human	ORF

389	28	73.7	107	5	ABP09313	Abp09313 Human ORF
390	28	73.7	108	5	ABP63769	ABP63769 Human ORF
391	28	73.7	111	3	AA645463	AA645463 Arabidops
392	28	73.7	120	2	AA112283	AA112283 Human 5'
393	28	73.7	124	8	AD126137	AD126137 Human pro
394	28	73.7	127	3	AA173436	AA173436 Human sec
395	28	73.7	130	3	AA600184	AA600184 Human sec
396	28	73.7	137	3	AA645462	AA645462 Arabidops
397	28	73.7	141	3	AA628194	AA628194 Arabidops
398	28	73.7	146	9	ABE41704	ABE41704 L. pneumo
399	28	73.7	153	9	ABE16602	ABE16602 B. subtil
400	28	73.7	157	7	ADC95874	ADC95874 E. faeciu
401	28	73.7	161	6	ABM69553	ABM69553 Photothab
402	28	73.7	162	4	AA698902	AA698902 E. coli G
403	28	73.7	167	8	ADY11547	ADY11547 Plant ful
404	28	73.7	169	8	ADY08496	ADY08496 Plant ful
405	28	73.7	178	8	ADN12271	ADN12271 TL-1R/TLR
406	28	73.7	180	6	ABR43184	ABR43184 Human REM
407	28	73.7	196	3	AA628193	AA628193 Arabidops
408	28	73.7	198	4	ABG21897	ABG21897 Novel hum
409	28	73.7	204	8	ADP56658	ADP56658 Human Tol
410	28	73.7	204	8	ADP48599	ADP48599 Human Tol
411	28	73.7	211	3	ABE38444	ABE38444 L. pneumo
412	28	73.7	214	3	AA643064	AA643064 Arabidops
413	28	73.7	214	3	AA618091	AA618091 Arabidops
414	28	73.7	214	4	AA692865	AA692865 C. glutami
415	28	73.7	214	7	ADL65953	ADL65953 C. glutam
416	28	73.7	214	7	ADL65575	ADL65575 C. glutam
417	28	73.7	214	9	ABD46839	ABD46839 Membrane
418	28	73.7	214	9	ABD72620	ABD72620 Corynebac
419	28	73.7	219	3	AA688060	AA688060 Human Tol
420	28	73.7	223	3	ABM97246	ABM97246 M. xanthu
421	28	73.7	230	3	AA628192	AA628192 Arabidops
422	28	73.7	237	7	ABO70264	ABO70264 Pseudomon
423	28	73.7	242	3	AA630771	AA630771 Arabidops
424	28	73.7	245	2	AA142302	AA142302 Trehalose
425	28	73.7	246	2	AA137559	AA137559 C. tracho
426	28	73.7	247	6	ADA34361	ADA34361 Acinetoba
427	28	73.7	253	9	ADY52914	ADY52914 Nostoc pu
428	28	73.7	254	3	ADY52966	ADY52966 Nostoc pu
429	28	73.7	257	3	AA628191	AA628191 Arabidops
430	28	73.7	259	4	ABG16894	ABG16894 Novel hum
431	28	73.7	261	8	ADT57199	ADT57199 Plant pol
432	28	73.7	267	2	AAW22303	AAW22303 Rat CRTI.
433	28	73.7	267	8	ADL27274	ADL27274 Amino aci
434	28	73.7	267	8	ABD40557	ABD40557 Murine XB
435	28	73.7	268	4	AA664860	AA664860 Heart mus
436	28	73.7	268	4	AA699931	AA699931 Human Mes
437	28	73.7	268	5	ABG61851	ABG61851 Prostate
438	28	73.7	268	9	ABE53855	ABE53855 Prostate
439	28	73.7	275	4	ABE66537	ABE66537 Drosophil
440	28	73.7	279	8	ADS22855	ADS22855 Bacterial
441	28	73.7	280	8	ADP29737	ADP29737 Human sec
442	28	73.7	281	5	AA630770	AA630770 Arabidops
443	28	73.7	288	5	ABE84985	ABE84985 Human PRO
444	28	73.7	288	5	ABE95591	ABE95591 Human ang
445	28	73.7	288	7	ADC60852	ADC60852 Human sto
446	28	73.7	288	7	ADD10627	ADD10627 Human sec
447	28	73.7	288	7	ADD11587	ADD11587 Human sec
448	28	73.7	288	7	ADD37380	ADD37380 Human sec
449	28	73.7	288	8	ABE41588	ABE41588 Human sec
450	28	73.7	288	8	ADH43771	ADH43771 Human PRO
451	28	73.7	288	8	ADK83116	ADK83116 Human PRO
452	28	73.7	288	8	ADN16966	ADN16966 Resistin
453	28	73.7	288	8	ADN97418	ADN97418 Human sto
454	28	73.7	288	9	ADY19786	ADY19786 PRO polyp
455	28	73.7	288	9	ADY14388	ADY14388 PRO polyp
456	28	73.7	288	9	ABE38715	ABE38715 L. pneumo
457	28	73.7	288	9	ABE40801	ABE40801 L. pneumo
458	28	73.7	288	9	ABE41938	ABE41938 L. pneumo
459	28	73.7	288	9	ABE40132	ABE40132 L. pneumo
460	28	73.7	290	3	AA620023	AA620023 Arabidops
461	28	73.7	290	3	AA654068	AA654068 Arabidops

462	28	73.7	291	8	ADO29637	ADO29637 Mouse GPC
463	28	73.7	294	9	ABE37487	ABE37487 L. pneumo
464	28	73.7	294	9	ABE36748	ABE36748 L. pneumo
465	28	73.7	297	3	AA620022	AA620022 Arabidops
466	28	73.7	297	3	AA640657	AA640657 Arabidops
467	28	73.7	302	3	AA618090	AA618090 Arabidops
468	28	73.7	302	3	AA643063	AA643063 Arabidops
469	28	73.7	304	4	ABG02510	ABG02510 Novel hum
470	28	73.7	304	10	AEF29059	AEF29059 Lead Cere
471	28	73.7	305	3	AA630769	AA630769 Arabidops
472	28	73.7	305	5	ABE90870	ABE90870 Herbicida
473	28	73.7	305	5	ABE84448	ABE84448 Slah-rela
474	28	73.7	305	6	ABP71511	ABP71511 Arabidops
475	28	73.7	305	8	ADQ74828	ADQ74828 Arabidops
476	28	73.7	305	8	ADU24241	ADU24241 Arabidops
477	28	73.7	305	9	ADX84931	ADX84931 Arabidops
478	28	73.7	306	6	ABP71512	ABP71512 Arabidops
479	28	73.7	307	3	AA651216	AA651216 Arabidops
480	28	73.7	307	9	ADX84933	ADX84933 Arabidops
481	28	73.7	308	3	AA654066	AA654066 Arabidops
482	28	73.7	308	3	AA620021	AA620021 Arabidops
483	28	73.7	308	3	AA609499	AA609499 Arabidops
484	28	73.7	309	3	AAE58153	AAE58153 Lung canc
485	28	73.7	311	3	AA630559	AA630559 Arabidops
486	28	73.7	311	5	ABE91960	ABE91960 Herbicida
487	28	73.7	315	5	ABE84458	ABE84458 Slah-rela
488	28	73.7	315	5	ABE84457	ABE84457 Slah-rela
489	28	73.7	317	6	ABP57736	ABP57736 S. pombe
490	28	73.7	317	6	ADX81925	ADX81925 Plant ful
491	28	73.7	320	8	ADP25091	ADP25091 PRO polyp
492	28	73.7	321	3	AA643062	AA643062 Arabidops
493	28	73.7	322	3	AA618089	AA618089 Arabidops
494	28	73.7	322	7	ABO80187	ABO80187 Pseudomon
495	28	73.7	322	8	ADX67518	ADX67518 Plant ful
496	28	73.7	325	3	AA640192	AA640192 Arabidops
497	28	73.7	325	3	AA620782	AA620782 Arabidops
498	28	73.7	326	3	AA651215	AA651215 Arabidops
499	28	73.7	326	5	ABE47476	ABE47476 Listeria
500	28	73.7	326	6	ABU31154	ABU31154 Protein e

ALIGNMENTS

RESULT 1						
ID	ADQ90197	standard; peptide; 7 AA.				
XX	ADQ90197;					
AC						
DT	21-OCT-2004	(first entry)				
XX						
DE	WMLSANS domain motoneuronotrophic factor peptide analogue.					
XX						
KW	motoneuronotrophic; motor neuron; Neurotrophic; Neuroprotective; Vulnerary;					
KM	Neuropeptide-Agonist; axon degeneration; muscle reinnervation;					
KW	peripheral nerve; neurodegenerative; wound healing.					
XX						
OS	Unidentified.					
XX						
FN	W02004065410-A2.					
PD	05-AUG-2004.					
XX						
PF	21-JAN-2004; 2004WO-US001468.					
XX						
PR	21-JAN-2003; 2003US-0441772P.					
XX						
PA	(GENE-) GENEVRON BIOPHARMACEUTICALS LLC.					
XX						
PI	Chau RMW, Ko TD;					
XX						
DR	WPI; 2004-562147/54.					

XX New motoneuronotrophic factor peptide analogue enhancing the viability of
 PT motor neurons, useful for target muscle reinnervation, treating
 PT peripheral nerve injuries or neurodegenerative diseases and in wound
 PT healing.
 XX
 PS Claim 1; SEQ ID NO 3; 40pp; English.
 XX
 CC The present invention relates to motoneuronotrophic factor peptide
 CC analogue, where the peptide analogue enhances the viability of motor
 CC neurons. The methods and compositions of the present invention are useful
 CC for promoting motor neuron viability and axon degeneration, target muscle
 CC reinnervation, treating peripheral nerve injuries, treating
 CC neurodegenerative disease and in wound healing. The present sequence
 CC represents WMLSAFS domain motoneuronotrophic factor peptide analogue.
 XX
 SQ Sequence 7 AA;
 Query Match 100.0%; Score 38; DB 8; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2.1e+06; Mismatches 0; Gaps 0;
 Matches 7; Conservative 0; Indels 0; Gaps 0;
 QY 1 WMLSAFS 7
 |||||
 Db 1 WMLSAFS 7
 RESULT 2
 ADQ90200
 ID ADQ90200 standard; peptide; 13 AA.
 XX
 AC ADQ90200;
 XX
 DT 21-OCT-2004 (first entry)
 XX
 DE MNTF 13 mer motoneuronotrophic factor peptide analogue.
 XX
 DE motoneuronotrophic; motor neuron; Neurotrophic; Neuroprotective; Vulnerary;
 KW Neuropeptide-Agonist; axon degeneration; muscle reinnervation;
 KW peripheral nerve; neurodegenerative; wound healing.
 XX
 OS Unidentified.
 XX
 PN WO2004065410-A2.
 XX
 PD 05-AUG-2004.
 XX
 PF 21-JAN-2004; 2004WO-US001468.
 XX
 PR 21-JAN-2003; 2003US-0441772P.
 XX
 PA (GENE-) GENERVON BIOPHARMACEUTICALS LLC.
 XX
 PI Chau RMW, Ko TD;
 XX
 DR WPI; 2004-562147/54.
 XX
 DE New motoneuronotrophic factor peptide analogue enhancing the viability of
 PT motor neurons, useful for target muscle reinnervation, treating
 PT peripheral nerve injuries or neurodegenerative diseases and in wound
 PT healing.
 XX
 PS Claim 7; SEQ ID NO 6; 40pp; English.
 XX
 CC The present invention relates to motoneuronotrophic factor peptide
 CC analogue, where the peptide analogue enhances the viability of motor
 CC neurons. The methods and compositions of the present invention are useful
 CC for promoting motor neuron viability and axon degeneration, target muscle
 CC reinnervation, treating peripheral nerve injuries, treating
 CC neurodegenerative disease and in wound healing. The present sequence
 CC represents a MNTF motoneuronotrophic factor peptide analogue.
 XX
 SQ Sequence 13 AA;

Query Match 100.0%; Score 38; DB 8; Length 13;
 Best Local Similarity 100.0%; Pred. No. 1; Mismatches 0; Gaps 0;
 Matches 7; Conservative 0; Indels 0; Gaps 0;
 QY 1 WMLSAFS 7
 |||||
 Db 2 WMLSAFS 8
 RESULT 3
 AAM59046
 ID AAM59046 standard; protein; 33 AA.
 XX
 AC AAM59046;
 XX
 DT 11-AUG-1998 (first entry)
 XX
 DE Human MNTF1-F6 protein fragment.
 XX
 DE Motoneuronotrophic factor; MNTF-1; MNTF1-F6; human; axon regeneration;
 KW motoneuron; diagnose; treatment; disease; wound healing; scar tissue;
 KW keloid.
 XX
 OS Homo sapiens.
 XX
 PN WO9813492-A2.
 XX
 PD 02-APR-1998.
 XX
 PF 22-SEP-1997; 97WO-US017142.
 XX
 PR 27-SEP-1996; 96US-0026792P.
 PR 15-NOV-1996; 96US-00751225.
 PR 12-SEP-1997; 97US-00928862.
 XX
 PA (KMBI-) KM BIOTECH INC.
 XX
 PI Chau RMW;
 XX
 DR WPI; 1998-230703/20.
 DR N-PSDB; AAV11748.
 XX
 PT Motoneuronotrophic factor MNTF1-F3 and MNTF1-F6 - useful for motoneuron
 PT regeneration, diagnosing or treating motoneuron disease and to accelerate
 PT wound healing without scar formation.
 XX
 PS Claim 4; Fig 2B; 78pp; English.
 XX
 CC This sequence, represents a fragment of a novel human motoneuronotrophic
 CC factor, MNTF1-F6. Such factors are used to promote regeneration of the
 CC axon of a motoneurone, to diagnose and treat motoneurone disease in a
 CC mammal or to accelerate wound healing whilst concomitantly minimising or
 CC inhibiting scar tissue and/or keloid formation in an area associated with
 CC a wound. For promoting axonal regeneration, the polypeptide is
 CC administered at a concentration of 5 ng-50 mg, whereas for inhibiting
 CC hereditary motoneurone disease, the dosage is 5-100 (especially 30-50)ng
 CC per kg body weight
 XX
 SQ Sequence 33 AA;
 Query Match 100.0%; Score 38; DB 2; Length 33;
 Best Local Similarity 100.0%; Pred. No. 2.8; Mismatches 0; Gaps 0;
 Matches 7; Conservative 0; Indels 0; Gaps 0;
 QY 1 WMLSAFS 7
 |||||
 Db 12 WMLSAFS 18
 RESULT 4
 AAO29914
 ID AAO29914 standard; protein; 33 AA.

XX AA029914;
 AC 03-SEP-2003 (first entry)
 DT
 XX
 DE Human motoneuronotrophic factor (MNTF)-F6 protein.
 XX
 XX Human; motoneuronotrophic factor; MNTF; therapy; radiation therapy;
 KM peripheral nerve injury; musculoskeletal disorder; spinal cord injury;
 KM head injury; stroke; neuromuscular degenerative disease; nootropic;
 KM amyotrophic lateral sclerosis; peripheral neuropathy; neuroprotective;
 KM muscular dystrophy; AIDS; spinal muscular atrophy; multiple sclerosis;
 KM scar; myasthenia gravis; sensory neuronal function disorder; vulnary;
 cancer.
 KW
 XX Homo sapiens.
 OS
 XX WO2003044175-A2.
 PN
 XX 30-MAY-2003.
 PD
 XX 19-NOV-2002; 2002WO-US037191.
 PF
 XX 20-NOV-2001; 2001US-00989481.
 PR
 XX (GENE-) GENERVON BIOPHARMACEUTICALS LLC.
 PA
 XX Chau RMW;
 PI
 XX WPI; 2003-457607/43.
 DR
 XX N-PSDB; AAL60573.
 DX
 XX Promoting the survival, growth, proliferation or maintenance of mammalian
 PT neurons by administering motoneuronotrophic factors, useful for treating
 PT musculoskeletal and neurodegenerative disorders and spinal cord injuries.
 PT
 XX Claim 1; Fig 2B; 90pp, English.
 PS
 XX The invention relates to a method for promoting the survival, growth,
 CC proliferation or maintenance of mammalian neurons by administering
 CC motoneuronotrophic factors (MNTF). The method is useful for treating
 CC peripheral nerve injuries, musculoskeletal disorders, spinal cord
 CC injuries, head injuries, strokes, neuromuscular degenerative diseases,
 CC amyotrophic lateral sclerosis, spinal muscular atrophy, peripheral
 CC neuropathy, diabetic peripheral neuropathy, peripheral neuropathy
 CC resulting from AIDS or radiation therapy for cancer, multiple sclerosis,
 CC muscular dystrophy, inhibition of scar formation, myasthenia gravis and
 CC sensory neuronal function disorders. The present sequence is human MNTF-
 CC F6 protein used to illustrate the method of the invention
 CC
 XX Sequence 33 AA;
 SQ
 Query Match 100.0%; Score 38; DB 6; Length 33;
 Best Local Similarity 100.0%; Pred. No. 2.8;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 WMLSAFS 7
 Db 12 WMLSAFS 18
 RESULT 5
 AD090195
 ID AD090195 standard; peptide; 33 AA.
 XX
 AC AD090195;
 XX
 DT 21-OCT-2004 (first entry)
 XX
 DE MNTF1-F6 33 mer motoneuronotrophic factor peptide analogue.
 XX
 XX motoneuronotrophic; motor neuron; Nootropic; Neuroprotective; Vulnary;
 KM Neuropeptide-Agonist; axon degeneration; muscle reinnervation;

KM peripheral nerve; neurodegenerative; wound healing.
 XX
 XX Undefined.
 OS
 XX WO2004065410-A2.
 PN
 XX 05-AUG-2004.
 PD
 XX 21-JAN-2004; 2004WO-US001468.
 PF
 XX 21-JAN-2003; 2003US-0441772P.
 PR
 XX (GENE-) GENERVON BIOPHARMACEUTICALS LLC.
 PA
 XX Chau RMW, Ko TD;
 PI
 XX WPI; 2004-562147/54.
 DR
 XX New motoneuronotrophic factor peptide analogue enhancing the viability of
 PT motor neurons, useful for target muscle reinnervation, treating
 PT peripheral nerve injuries or neurodegenerative diseases and in wound
 PT healing.
 PT
 XX Claim 2; SEQ ID NO 1; 40pp; English.
 PS
 XX The present invention relates to motoneuronotrophic factor peptide
 CC analogue, where the peptide analogue enhances the viability of motor
 CC neurons. The methods and compositions of the present invention are useful
 CC for promoting motor neuron viability and axon degeneration, target muscle
 CC reinnervation, treating peripheral nerve injuries, treating
 CC neurodegenerative disease and in wound healing. The present sequence
 CC represents MNTF1-F6 33 mer motoneuronotrophic factor peptide analogue.
 CC
 XX Sequence 33 AA;
 SQ
 Query Match 100.0%; Score 38; DB 8; Length 33;
 Best Local Similarity 100.0%; Pred. No. 2.8;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 WMLSAFS 7
 Db 12 WMLSAFS 18
 RESULT 6
 ABB63367
 ID ABB63367 standard; protein; 1760 AA.
 XX
 AC ABB63367;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 16893.
 XX
 KM Drosophila; developmental biology; cell signalling; insecticide;
 KM pharmaceutical.
 XX
 XX Drosophila melanogaster.
 OS
 XX WO200171042-A2.
 PN
 XX 27-SEP-2001.
 PD
 XX 23-MAR-2001; 2001WO-US009231.
 PF
 XX 23-MAR-2000; 2000US-0191637P.
 PR
 XX 11-JUL-2000; 2000US-00614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX WPI; 2001-656860/75.
 DR

DR N-PSDB; ABL07470.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Disclosure; SEQ ID NO 16893; 21bp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins (AB57737-
CC AB572072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1760 AA;
XX
Query Match 92.1%; Score 35; DB 4; Length 1760;
Best Local Similarity 85.7%; Pred. No. 7.7e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 WMSAFS 7
DB 966 MWVSARS 972
XX
RESULT 7
ABG70020
ID ABG70020 standard; protein; 1760 AA.
XX
AC ABG70020;
XX
DT 05-NOV-2002 (first entry)
XX
DE Larval viability associated protein #19.
XX
XX Fruit fly; larval viability; insecticidal activity; maize; wheat; oat;
KM rye; sorghum; rice; barley; millet; turf; cotton; sugarcane; sugar beet;
KM oilseed rape; soybean; vegetable crop; fruit.
XX
XX Drosophila melanogaster.
XX
PN WO200257455-A2.
XX
PD 25-JUL-2002.
XX
PF 18-JAN-2002; 2002WO-US001568.
XX
PR 18-JAN-2001; 2001US-0262351P.
XX
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Stam L, Bachmann J, Broadus J, Kamdar KP;
XX
XX WPI; 2002-590746/63.
DR N-PSDB; AB551396.
XX
XX Identifying inhibitors of activity of proteins essential for Drosophila
PT larval viability comprising expressing in a host a protein essential for
PT larval activity and identifying compounds that inhibit or interact with
PT the protein.
XX
PS Claim 1; Page 120-126; 169pp; English.
XX
CC The invention describes a method of identifying compounds that inhibit
CC the activity of, or that interact with a protein essential for Drosophila
CC larval viability comprising expressing in a recombinant host a DNA
CC molecule to produce a protein essential for larval viability. The method
CC is useful for identifying compounds with insecticidal activity. Compounds

CC identified are useful as insecticides in crops such as maize, wheat,
CC oats, rye, sorghum, rice, barley, millet, turf, cotton, sugarcane, sugar
CC beet, oilseed rape, soybeans, vegetable crops and fruits. This is the
CC amino acid sequence of a fruit fly larval viability associated protein
XX
SQ Sequence 1760 AA;
XX
Query Match 92.1%; Score 35; DB 5; Length 1760;
Best Local Similarity 85.7%; Pred. No. 7.7e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 WMSAFS 7
DB 966 MWVSARS 972
XX
RESULT 8
ADX56419
ID ADX56419 standard; peptide; 17 AA.
XX
AC ADX56419;
XX
DT 05-MAY-2005 (first entry)
XX
DE Cardiovascular disorder plasma protein tryptic fragment SEQ ID NO 236.
XX
XX coronary artery disease; primary biliary cirrhosis; gallstones;
KM celiac disease; irritable bowel syndrome; diabetes; scleroderma; nausea;
KM emesis; constipation; diarrhea; cardiovascular disease; immune disorder;
KM endocrine disease; gastrointestinal disease; metabolic disorder;
KM dermatological disease; musculoskeletal disease; Cardiant; Vasotropic;
KM Hepatotropic; Litholytic; Immunosuppressive; Gastrointestinal.Gen.;
KM Antiinflammatory; Antidiabetic; Dermatological; Antiemetic; Laxative;
KM Antidiarrheic; diagnosis.
XX
XX Homo sapiens.
XX
PN WO2005015206-A2.
XX
PD 17-FEB-2005.
XX
PF 06-AUG-2004; 2004WO-EP008860.
XX
PR 08-AUG-2003; 2003US-0493599P.
PR 08-AUG-2003; 2003US-0493836P.
PR 08-AUG-2003; 2003US-0493867P.
PR 08-AUG-2003; 2003US-0493985P.
XX
PA (XENO-) XENOVA LTD.
PA (NOVS) NOVARTIS AG.
PA (NOVS) NOVARTIS PHARMA GMBH.
XX
PI Argoud-Puy G, Bedert N, Bougueleret L, Cusin I, Mahe E;
PI Niknejad A, Refas S, Rose K, Saudrais C, Scherer A, Papoian R;
XX
XX WPI; 2005-195824/20.
DR
XX
XX Screening and/or diagnosing cardiovascular disorder in subject involves
PT detecting and/or quantifying level of polypeptide in biological sample
PT from subject and comparing with control sample.
XX
PS Claim 1; SEQ ID NO 236; 349pp; English.
XX
XX The invention relates to a method of screening and/or diagnosing a
CC cardiovascular disorder (CD) in a subject which comprises detecting
CC and/or quantifying the level of a polypeptide in a biological sample from
CC the subject and comparing the level to that of control sample. The method
CC is useful for screening, diagnosing and treating coronary artery disease,
CC biliary cirrhosis, gallstones, celiac disease, irritable bowel syndrome,
CC diabetes, scleroderma, nausea, vomiting, constipation and diarrhea. The
CC method is rapid and efficient. The present sequence represents a
CC cardiovascular disorder plasma protein tryptic fragment.
XX

SQ Sequence 17 AA;

Query Match 86.8%; Score 33; DB 9; Length 17;

Best Local Similarity 85.7%; Pred. No. 13;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7

Db 5 WMLSPFS 11

RESULT 9

AAB29781 ID AAB29781 standard; peptide; 35 AA.

AC AAB29781;

XX 28-FEB-2001 (first entry)

DE Human MSF-derived tribonectin fragment #3.

XX Human tribonectin; MSF, megakaryocyte stimulating factor;

KM alternative splicing; joint boundary lubricant; O-linked oligosaccharide;

KM osteoarthritis; tribosupplementation; tissue adhesion inhibition;

KM friction coefficient reduction; gene therapy; antiarthritic; osteopathic.

OS Homo sapiens.

XX WO200064930-A2.

PN 02-NOV-2000.

XX 24-APR-2000; 2000WO-US010953.

XX 23-APR-1999; 99US-00298970.

XX (RHOD-) RHODE ISLAND HOSPITAL LIFESPAN PARTNER.

XX Jay GD;

XX WPI; 2001-024673/03.

XX Novel tribonectin polypeptide useful as lubricant for treating

XX osteoarthritis, comprises O-linked lubricating moiety.

XX Disclosure; Fig 1; 47pp; English.

XX The invention relates to a human tribonectin which is a product of

XX alternative splicing of the human MSF (megakaryocyte stimulating factor)

XX gene. The tribonectin has at least one O-linked oligosaccharide

XX lubricating moiety and has a polypeptide sequence comprising 1-76 repeats

XX of a motif having at least 50% identity to the sequence KEPTPT

XX (AAB29781). The invention also relates to a nucleic acid encoding a human

XX MSF-derived tribonectin; a biocompatible composition comprising a human

XX tribonectin for inhibiting tissue adhesion formation; and a method of

XX diagnosing osteoarthritis or a predisposition to osteoarthritis by

XX measuring the amount of MSF or its fragment in a biological sample of a

XX mammal, wherein an increased amount of MSF compared to a control

XX indicates the presence of or predisposition to developing osteoarthritis.

XX The tribonectin and DNA encoding it are useful in the treatment of

XX osteoarthritis, where they may be used for lubricating mammalian joints,

XX such as articulating joints of humans, dogs or horses. The tribonectin,

XX when formulated as a membrane, foam, gel or fibre, is useful for

XX inhibiting adhesion between two surfaces such as the injured tissues of a

XX mammal, where the injury is caused by a surgical insertion or trauma, or of

XX an artificial device e.g., an orthopaedic implant. In particular, one of

XX the surfaces is pericardial tissue. DNA encoding a tribonectin may be

XX used in gene therapy. The present sequence represents a fragment of a

XX human MSF-derived tribonectin

XX Sequence 35 AA;

SQ Query Match 86.8%; Score 33; DB 4; Length 35;

Best Local Similarity 85.7%; Pred. No. 29;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7

Db 6 WMLSPFS 12

RESULT 10

AAB44007 ID AAB44007 standard; protein; 61 AA.

AC AAB44007;

XX 08-FEB-2001 (first entry)

DE Human cancer associated protein sequence SEQ ID NO:1452.

XX Human, cancer associated gene; cancer antigen; detection; cancer;

KM diagnosis; cytostatic; proliferative; vulnery; immunomodulator;

KM antidiabetic; antisthmatic; antineumatic; antiarthritic; antiviral;

KM antiinflammatory; antihypoid; antiallergic; antibacterial; cardiac;

KM dermatological; neuroprotective; thrombolytic; coagulant; neotropic;

KM vasotropic; antipsoriatic; angiogenic; gene therapy; inflammation;

KM immune disorder; haematopoietic cell disorder; autoimmune disorder;

KM allergic reaction; graft versus host disease; organ rejection;

KM haemostatic; thrombolytic; cardiovascular disorder; infection;

KM neurological disease; drug screening.

XX Homo sapiens.

XX WO200055350-A1.

XX 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US005882.

XX 12-MAR-1999; 99US-0124270P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;

XX WPI; 2000-587533/55.

XX N-PSDB; AAC78216.

XX Novel isolated nucleic acids comprising sequences encoding peptides

XX useful for treating or diagnosing e.g. cancer.

XX Claim 11; Page 2133; 2352pp; English.

XX AAC77607 to AAC78448 encode the human cancer associated proteins given in

XX AAB43398 to AAB44239. The proteins can have activities based on the

XX tissues and cells the genes are expressed in. Example of activities

XX include: cytostatic; proliferative; vulnery; immunomodulator;

XX antidiabetic; antisthmatic; antineumatic; antiarthritic;

XX antiinflammatory; antihypoid; antiallergic; antibacterial; antiviral;

XX dermatological; neuroprotective; cardiac; thrombolytic; coagulant;

XX neotropic; vasotropic; antipsoriatic and angiogenic. The

XX polynucleotides and polypeptides can be used for preventing, treating or

XX ameliorating medical conditions and diagnosing pathological conditions.

XX Polynucleotides, polypeptides, antibodies, agonists and antagonists from

XX the present invention may be used to treat immune disorders by activating

XX or inhibiting the proliferation, differentiation or mobilisation of

XX immune cells, to treat disorders of haematopoietic cells, autoimmune

XX disorders, allergic reactions, graft versus host disease and organ

XX rejection, modulate haemostatic or thrombolytic activity, modulate

XX inflammation, cancers, cardiovascular disorders, neurological disease and

XX bacterial or viral infections. The peptides, nucleotides, antibodies,

XX agonists and antagonists may be also be used in drug screens. AAC78449 to

XX AAC78457 and AAB44240 represent sequences used in the exemplification of

XX the present invention

SQ Sequence 61 AA;

Query Match 86.8%; Score 33; DB 3; Length 61;
 Best Local Similarity 85.7%; Pred. No. 52;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
 Db 37 WMLSPFS 43

RESULT 11
 ID AAO18834 standard; protein; 538 AA.

XX AAO18834;
 AC
 XX
 XX
 DT 29-OCT-2002 (first entry)
 XX
 XX 3' cartilage superficial zone protein coding sequence encoded protein.
 DE
 XX SZP; superficial zone protein; cartilage; lubrication; human;
 KW degenerative joint condition; arthritis; osteoporosis; trauma; CACP;
 KW chondroitin sulphate substitution consensus; antiarthritic;
 KW antirheumatic; osteopathic; antigout; antiinflammatory; dermatological;
 KW immunosuppressive.
 XX
 XX Homo sapiens.
 OS
 XX
 XX Key Location/Qualifiers
 FT Misc-difference 114 /note="encoded by ACTACT"
 FT
 XX WO200262847-A2.
 PN
 XX 15-AUG-2002.
 PD
 XX 31-DEC-2001; 2001WO-US050379.
 PF
 XX 29-DEC-2000; 2000US-0258920P.
 PR
 XX
 XX (GLAX) GLAXO GROUP LTD.
 PA (RUSH-) RUSH PRESBYTERIAN ST LUKE MEDICAL CENT.
 XX
 PI Hutchins JT, Kuettner KE, Schmid TM, Schumacher BL, Su J;
 PI Dixon EP;
 XX
 XX WPI; 2002-636585/68.
 DR N-PSDB; AAL49079.
 XX
 XX New purified superficial zone protein (SZP) polypeptides, useful for
 PT treating degenerative joint conditions, e.g. osteoarthritis, rheumatoid
 PT arthritis, gout, spondylarthritis, synovitis, tendonitis, lupus, or
 PT osteoporosis.
 XX
 XX Claim 59; Page 86-87; 89pp; English.
 PS
 XX The present invention provides the protein and coding sequences of human
 CC superficial zone protein (SZP). The protein is involved in the
 CC lubrication of joints, and the sequences can be used in the treatment of
 CC degenerative joint conditions or to delay symptoms of a degenerative
 CC joint condition, e.g. osteoarthritis, rheumatoid arthritis, gout,
 CC psoriatic arthritis, reactive arthritis, viral or post viral arthritis,
 CC spondylarthritis, juvenile arthritis, synovitis, tendonitis, systemic
 CC lupus erythematosus, CACP, osteoporosis or trauma. The present sequence
 CC is the protein encoded by the human 3' cartilage SZP cDNA
 XX
 XX Sequence 538 AA;

Query Match 86.8%; Score 33; DB 5; Length 538;
 Best Local Similarity 85.7%; Pred. No. 5.4e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
 Db 305 WMLSPFS 311

RESULT 12
 ID AAB29778 standard; protein; 902 AA.
 XX AAB29778;
 AC
 XX
 XX 28-FEB-2001 (first entry)
 DT
 XX
 XX Human MSF-derived tribonectin.
 DE
 XX Human tribonectin; MSF; megakaryocyte stimulating factor;
 KW alternative splicing; joint boundary lubricant; O-linked oligosaccharide;
 KW osteoarthritis; tribosupplementation; tissue adhesion inhibition;
 KW friction coefficient reduction; gene therapy; antiarthritic; osteopathic.
 KW
 XX Homo sapiens.
 OS
 XX WO200064930-A2.
 PN
 XX 02-NOV-2000.
 PD
 XX 24-APR-2000; 2000WO-US010953.
 PF
 XX 23-APR-1999; 99US-00298970.
 PR
 XX (RHOD-) RHODE ISLAND HOSPITAL LIFESPAN PARTNER.
 PA
 XX Jay GD;
 PI
 XX WPI; 2001-024673/03.
 DR
 XX
 XX Novel tribonectin polypeptide useful as lubricant for treating
 PT osteoarthritis, comprises O-linked lubricating moiety.
 PT
 XX Disclosure; Fig 1; 47pp; English.
 PS
 XX The invention relates to a human tribonectin which is a product of
 CC alternative splicing of the human MSF (megakaryocyte stimulating factor)
 CC gene. The tribonectin has at least one O-linked oligosaccharide
 CC lubricating moiety and has a polypeptide sequence comprising 1-76 repeats
 CC of a motif having at least 50% identity to the sequence KEDAPPT
 CC (AAB29774). The invention also relates to a nucleic acid encoding a human
 CC MSF-derived tribonectin; a biocompatible composition comprising a human
 CC tribonectin for inhibiting tissue adhesion formation; and a method of
 CC diagnosing osteoarthritis or a predisposition to osteoarthritis by
 CC measuring the amount of MSF or its fragment in a biological sample of a
 CC mammal, wherein an increased amount of MSF compared to a control
 CC indicates the presence of or predisposition to developing osteoarthritis.
 CC The tribonectin and DNA encoding it are useful in the treatment of
 CC osteoarthritis, where they may be used for lubricating mammalian joints,
 CC such as articulating joints of humans, dogs or horses. The tribonectin,
 CC when formulated as a membrane, foam, gel or fibre, is useful for
 CC inhibiting adhesion between two surfaces such as the injured tissues of a
 CC mammal, where the injury is caused by a surgical incision or trauma, or
 CC an artificial device e.g., an orthopaedic implant. In particular, one of
 CC the surfaces is pericardial tissue. DNA encoding a tribonectin may be
 CC used in gene therapy. The present sequence represents a substantial
 CC portion of a human MSF-derived tribonectin
 XX
 XX Sequence 902 AA;

Query Match 86.8%; Score 33; DB 4; Length 902;
 Best Local Similarity 85.7%; Pred. No. 9.3e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
 Db 811 WMLSPFS 817

```

RESULT 13
AD067713
ID AD067713 standard; protein; 933 AA.
XX
AC AD067713;
XX
DT 07-OCT-2004 (first entry)
XX
DE Novel human protein sequence #2379.
XX
KM osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;
KW gene therapy; diagnostic marker; morbid state; osteoporosis;
KM neurological disease; Alzheimer's disease; Parkinson's disease; dementia;
KM cancer.
XX
OS Homo sapiens.
XX
PN EP1440981-A2.
XX
PD 28-JUL-2004.
XX
PF 21-JAN-2004; 2004EP-00001196.
XX
PR 21-JAN-2003; 2003JP-00102206.
PR 09-MAY-2003; 2003JP-00131392.
XX
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Nagai K, Irie R;
XX
DR WPI; 2004-535376/52.
DR N-PSDB; AD067406.
XX
PT Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,
PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
XX
PS Claim 1; SEQ ID NO 4874; 2449pp; English.
XX
CC The invention relates to 2495 novel polynucleotides (I) and their encoded
CC polypeptides, sequences hybridizing to these nucleotides, sequences
CC encoding partial polypeptides and sequences having 70% or 90% identity to
CC the nucleotide and protein sequences. The nucleotides and polypeptides
CC are useful as diagnostic markers or therapeutic target for the diseases
CC or morbid states. They are also useful for treating osteoporosis,
CC neurological diseases, Alzheimer's diseases, Parkinson's diseases,
CC dementia and various cancers. This sequence corresponds to a protein
CC sequence of the invention.
XX
SQ Sequence 933 AA;

Query Match 86.8%; Score 33; DB 8; Length 933;
Best Local Similarity 85.7%; Pred. No. 9.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 WMLSPFS 7
Db 700 WMLSPFS 706

RESULT 14
AD203664
ID AD203664 standard; protein; 981 AA.
XX
AC AD203664;
XX
DT 02-JUN-2005 (first entry)
XX
PR PG4-Lub:1 protein.
XX
KW lubricin; antiarthritic; osteopathic; antirheumatic; cyrostatic;
KM

```

```

KM cardiant; respiratory-gen.; osteoarthritis; rheumatoid arthritis;
KW neoplasm; musculoskeletal disease; immune disorder; inflammation.
XX
OS Synthetic.
XX
PN WO2005016130-A2.
XX
PD 24-FEB-2005.
XX
PF 13-AUG-2004; 2004WO-US026508.
XX
PR 14-AUG-2003; 2003US-0495741P.
XX
PA (AMHP) WYETH.
XX
PI Flannery CR, Corcoran CJ, Freeman BA, Racie LA;
XX
DR WPI; 2005-272808/28.
DR N-PSDB; AD203663.
XX
PT Novel recombinant lubricin useful for treating osteoarthritis or
PT rheumatoid arthritis, or as lubricants, anti-adhesive agent or intra
PT articular supplements for synovial joints, meniscus, tendon, peritoneum,
PT pericardium or pleura.
XX
PS Claim 11; SEQ ID NO 7; 19pp; English.
XX
CC The invention relates to an isolated lubricin protein. A lubricin
CC containing composition is useful for treating a subject which involves
CC administering same to a tissue (e.g. cartilage, synovium, meniscus,
CC tendon, peritoneum, pericardium, or pleura (preferably cartilage)) of the
CC subject (e.g. mouse, rat, cat, dog, horse or human (preferably human)).
CC The above method further involves providing an anesthetic,
CC antiinflammatory drug and antibiotic to the subject, aspirating fluid
CC from the subject, washing tissue of the subject and imaging tissue of the
CC subject. The protein is useful for treating osteoarthritis or rheumatoid
CC arthritis, or as lubricants, anti-adhesive agent or intra articular
CC supplements for synovial joints, meniscus, tendon, peritoneum,
CC pericardium or pleura. An anti-lubricin antibody is useful for treating
CC cancer e.g. synovias. The present sequence represents the amino acid
CC sequence of the PG4-Lub:1 protein.
XX
SQ Sequence 981 AA;

Query Match 86.8%; Score 33; DB 9; Length 981;
Best Local Similarity 85.7%; Pred. No. 1e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 WMLSPFS 7
Db 748 WMLSPFS 754

RESULT 15
AD203668
ID AD203668 standard; protein; 1007 AA.
XX
AC AD203668;
XX
DT 02-JUN-2005 (first entry)
XX
PR PG4-Lub:2 protein.
XX
KW lubricin; antiarthritic; osteopathic; antirheumatic; cyrostatic;
KW cardiant; respiratory-gen.; osteoarthritis; rheumatoid arthritis;
KW neoplasm; musculoskeletal disease; immune disorder; inflammation.
XX
OS Synthetic.
XX
PN WO2005016130-A2.
XX
PD 24-FEB-2005.
XX

```

PF 13-AUG-2004; 2004MO-US026508.
 XX
 XX 14-AUG-2003; 2003US-0495741P.
 PR
 XX (AMHP) WYETH.
 PA
 XX Flannery CR, Corcoran CJ, Freeman BA, Racie LA;
 PI
 XX WPI; 2005-272808/28.
 DR
 XX Novel recombinant lubricin useful for treating osteoarthritis or
 PT rheumatoid arthritis, or as lubricants, anti-adhesive agent or intra
 PT articular supplements for synovial joints, meniscus, tendon, peritoneum,
 PT pericardium or pleura.
 PT
 XX Claim 11; SEQ ID NO 11; 19pp; English.
 PS
 XX The invention relates to an isolated lubricin protein. A lubricin
 CC containing composition is useful for treating a subject which involves
 CC administering same to a tissue (e.g. cartilage, synovium, meniscus,
 CC tendon, peritoneum, pericardium, or pleura (preferably cartilage)) of the
 CC subject (e.g. mouse, rat, cat, dog, horse or human (preferably human)).
 CC The above method further involves providing an anesthetic,
 CC antiinflammatory drug and antibiotic to the subject, aspirating fluid
 CC from the subject, washing tissue of the subject and imaging tissue of the
 CC subject. The protein is useful for treating osteoarthritis or rheumatoid
 CC arthritis, or as lubricants, anti-adhesive agent or intra articular
 CC supplements for synovial joints, meniscus, tendon, peritoneum,
 CC pericardium or pleura. An anti-lubricin antibody is useful for treating
 CC cancer e.g. synoviomias. The present sequence represents the amino acid
 CC sequence of the PRG4-Lub:2 protein.
 CC
 SQ Sequence 1007 AA;
 Query Match 86.8%; Score 33; DB 9; Length 1007;
 Best Local Similarity 85.7%; Pred. No. 1e+03; 1; Indels 0; Gaps 0;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 WMLSAFS 7
 Db 774 WMLSPFS 780
 RESULT 16
 AD203672
 ID AD203672 standard; protein; 1038 AA.
 XX
 AC AD203672;
 XX
 XX 02-JUN-2005 (first entry)
 DT
 XX
 DE PRG4-Lub:3 protein.
 XX
 XX lubricin; antiarthritic; osteopathic; antirheumatic; cytosratic;
 KW cardiac; respiratory-gen.; osteoarthritis; rheumatoid arthritis;
 KW neoplasm; musculoskeletal disease; immune disorder; inflammation.
 XX
 OS Synthetic.
 XX
 PN WO2005016130-A2.
 XX
 PD 24-FEB-2005.
 XX
 PF 13-AUG-2004; 2004MO-US026508.
 XX
 PR 14-AUG-2003; 2003US-0495741P.
 XX
 XX (AMHP) WYETH.
 PA
 XX Flannery CR, Corcoran CJ, Freeman BA, Racie LA;
 PI
 XX WPI; 2005-272808/28.
 DR N-PSDB; AD203671.

XX
 XX Novel recombinant lubricin useful for treating osteoarthritis or
 PT rheumatoid arthritis, or as lubricants, anti-adhesive agent or intra
 PT articular supplements for synovial joints, meniscus, tendon, peritoneum,
 PT pericardium or pleura.
 PT
 XX Claim 11; SEQ ID NO 15; 19pp; English.
 PS
 XX The invention relates to an isolated lubricin protein. A lubricin
 CC containing composition is useful for treating a subject which involves
 CC administering same to a tissue (e.g. cartilage, synovium, meniscus,
 CC tendon, peritoneum, pericardium, or pleura (preferably cartilage)) of the
 CC subject (e.g. mouse, rat, cat, dog, horse or human (preferably human)).
 CC The above method further involves providing an anesthetic,
 CC antiinflammatory drug and antibiotic to the subject, aspirating fluid
 CC from the subject, washing tissue of the subject and imaging tissue of the
 CC subject. The protein is useful for treating osteoarthritis or rheumatoid
 CC arthritis, or as lubricants, anti-adhesive agent or intra articular
 CC supplements for synovial joints, meniscus, tendon, peritoneum,
 CC pericardium or pleura. An anti-lubricin antibody is useful for treating
 CC cancer e.g. synoviomias. The present sequence represents the amino acid
 CC sequence of the PRG4-Lub:3 protein.
 CC
 SQ Sequence 1038 AA;
 Query Match 86.8%; Score 33; DB 9; Length 1038;
 Best Local Similarity 85.7%; Pred. No. 1.1e+03; 1; Indels 0; Gaps 0;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 WMLSAFS 7
 Db 805 WMLSPFS 811
 RESULT 17
 AD203676
 ID AD203676 standard; protein; 1069 AA.
 XX
 AC AD203676;
 XX
 XX 02-JUN-2005 (first entry)
 DT
 XX
 DE PRG4-Lub:4 protein.
 XX
 XX lubricin; antiarthritic; osteopathic; antirheumatic; cytosratic;
 KW cardiac; respiratory-gen.; osteoarthritis; rheumatoid arthritis;
 KW neoplasm; musculoskeletal disease; immune disorder; inflammation.
 XX
 OS Synthetic.
 XX
 PN WO2005016130-A2.
 XX
 PD 24-FEB-2005.
 XX
 PF 13-AUG-2004; 2004MO-US026508.
 XX
 PR 14-AUG-2003; 2003US-0495741P.
 XX
 XX (AMHP) WYETH.
 PA
 XX Flannery CR, Corcoran CJ, Freeman BA, Racie LA;
 PI
 XX WPI; 2005-272808/28.
 DR N-PSDB; AD203675.
 XX
 XX Novel recombinant lubricin useful for treating osteoarthritis or
 PT rheumatoid arthritis, or as lubricants, anti-adhesive agent or intra
 PT articular supplements for synovial joints, meniscus, tendon, peritoneum,
 PT pericardium or pleura.
 PT
 XX Claim 11; SEQ ID NO 19; 19pp; English.
 PS
 XX The invention relates to an isolated lubricin protein. A lubricin

CC containing composition is useful for treating a subject which involves
 CC administering same to a tissue (e.g. cartilage, synovium, meniscus,
 CC tendon, peritoneum, pericardium, or pleura (preferably cartilage)) of the
 CC subject (e.g. mouse, rat, cat, dog, horse or human (preferably human)).
 CC The above method further involves providing an anesthetic,
 CC antiinflammatory drug and antibiotic to the subject, aspirating fluid
 CC from the subject, washing tissue of the subject and imaging tissue of the
 CC subject. The protein is useful for treating osteoarthritis or rheumatoid
 CC arthritis, or as lubricants, anti-adhesive agent or intra articular
 CC supplements for synovial joints, meniscus, tendon, peritoneum,
 CC pericardium or pleura. An anti-lubricin antibody is useful for treating
 CC cancer e.g. synoviomias. The present sequence represents the amino acid
 CC sequence of the PRG4-Lub:5 protein.

XX Sequence 1069 AA;
 SQ

Query Match 86.8%; Score 33; DB 9; Length 1069;
 Best Local Similarity 85.7%; Pred. No. 1.1e+03;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
 |||||
 Db 836 WMLSPFS 842

RESULT 18
 AD203680
 ID AD203680 standard; protein; 1100 AA.
 XX
 AC AD203680;
 XX
 DT 02-JUN-2005 (first entry)
 XX
 DE PRG4-Lub:5 protein.
 XX
 XX Lubricin; antiarthritic; osteopathic; antirheumatic; cyrostatic;
 KW cardiant; respiratory-gen.; osteoarthritis; rheumatoid arthritis;
 KW neoplasm; musculoskeletal disease; immune disorder; inflammation.
 XX
 OS Synthetic.
 XX
 PN W02005016130-A2.
 XX
 PD 24-FEB-2005.
 XX
 PF 13-AUG-2004; 2004WO-US026508.
 XX
 PR 14-AUG-2003; 2003US-0495741P.
 XX
 PA (AMHP) WYETH.
 XX
 PI Flannery CR, Corcoran CJ, Freeman BA, Racie LA;
 XX
 DR WPI: 2005-272808/28.
 DR N-PSDB; AD203679.

PT Novel recombinant lubricin useful for treating osteoarthritis or
 PT rheumatoid arthritis, or as lubricants, anti-adhesive agent or intra
 PT articular supplements for synovial joints, meniscus, tendon, peritoneum,
 PT pericardium or pleura.
 XX
 PS Claim 11; SEQ ID NO 23; 19pp; English.
 XX
 CC The invention relates to an isolated lubricin protein. A lubricin
 CC containing composition is useful for treating a subject which involves
 CC administering same to a tissue (e.g. cartilage, synovium, meniscus,
 CC tendon, peritoneum, pericardium, or pleura (preferably cartilage)) of the
 CC subject (e.g. mouse, rat, cat, dog, horse or human (preferably human)).
 CC The above method further involves providing an anesthetic,
 CC antiinflammatory drug and antibiotic to the subject, aspirating fluid
 CC from the subject, washing tissue of the subject and imaging tissue of the
 CC subject. The protein is useful for treating osteoarthritis or rheumatoid
 CC arthritis, or as lubricants, anti-adhesive agent or intra articular

CC supplements for synovial joints, meniscus, tendon, peritoneum,
 CC pericardium or pleura. An anti-lubricin antibody is useful for treating
 CC cancer e.g. synoviomias. The present sequence represents the amino acid
 CC sequence of the PRG4-Lub:5 protein.

XX Sequence 1100 AA;
 SQ

Query Match 86.8%; Score 33; DB 9; Length 1100;
 Best Local Similarity 85.7%; Pred. No. 1.2e+03;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
 |||||
 Db 867 WMLSPFS 873

RESULT 19
 ADK67911
 ID ADK67911 standard; protein; 1270 AA.
 XX
 AC ADK67911;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE Human extracellular messenger (EXMES) polypeptide.
 XX
 XX Human; extracellular messenger; EXMES; respiratory-gen.; antiallergic;
 KW antiasthmatic; antiinflammatory; antidiabetic; neuroprotective;
 KW muscular-gen.; antiarthritic; osteopathic; hepatotropic; antipsoriatic;
 KW virucide; fungicide; antiparasitic; protozoacide; antihelmintic;
 KW cyrostatic; gene therapy.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..29
 FT /label=Signal_peptide
 FT /note="Spans residues 1 to 18, 20, 21, 24, 27 or 29
 FT according to identification method"
 XX
 PN W02004013292-A2.
 XX
 PD 12-FEB-2004.
 XX
 PF 30-JUL-2003; 2003WO-US024084.
 XX
 PR 02-AUG-2002; 2002US-0400810P.
 PR 19-SEP-2002; 2002US-0412197P.
 PR 04-OCT-2002; 2002US-0416004P.
 PR 08-NOV-2002; 2002US-0424862P.
 XX
 PA (INCY-) INCYTE CORP.
 XX
 PI Elliott VS, Khare R, Tran UK, Swarnakar A, Marguis JP;
 PI Richardson IW, Emerling BM, Lindquist EA, Chawla NK, Ramkumar J;
 PI Lee SY;
 XX
 DR WPI: 2004-157116/15.
 DR N-PSDB; ADK67916.

PT New extracellular messengers and nucleic acids, useful for diagnosing,
 PT treating or preventing e.g. allergies, anemia, Crohn's disease, diabetes
 PT mellitus, myasthenia gravis, osteoporosis, psoriasis, cancers or
 PT autoimmune thyroiditis.
 XX
 PS Claim 59; SEQ ID NO 4; 165pp; English.
 XX
 CC The present sequence is that of novel human extracellular messenger
 CC (EXMES) Incyte ID NO: 7513017CD1 polypeptide. The protein shows homology
 CC to human megakaryocyte stimulating factor. The invention provides EXMES
 CC polynucleotides and polypeptides, as well as expression vectors, host
 CC cells, antibodies, agonists and antagonists, and methods for diagnosing,
 CC treating or preventing disorders associated with aberrant expression of

CC EXMES, especially autoimmune and inflammatory disorders, cell
 CC proliferative disorders and endocrine disorders, e.g. adult respiratory
 CC distress syndrome, allergy, anaemia, erythroblastosis foetalis, Crohn's
 CC disease, diabetes mellitus, myasthenia gravis, osteoarthritis,
 CC osteoporosis, pancreatitis, psoriasis, viral, bacterial, fungal,
 CC parasitic, protozoal or helminthic infections, cancers, autoimmune
 CC thyroiditis, cretinism, Plummer's disease or thyroid carcinoma.
 CC Embodiments also provide methods for using the purified EXMES and/or
 CC their encoding polynucleotides for facilitating the drug discovery
 CC process, including determining of efficacy, dosage, toxicity and
 CC pharmacology, and for investigating the pathogenesis of diseases and
 CC medical conditions.

XX
 SQ Sequence 1270 AA;

Query Match 86.8%; Score 33; DB 8; Length 1270;
 Best Local Similarity 85.7%; Pred. No. 1.3e+03;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSPS 7
 Db 1037 WMLSPS 1043

RESULT 20
 ID AAM24322 standard; protein; 1299 AA.
 AC AAM24322;
 XX
 DT 12-OCT-2001 (first entry)
 XX
 DE Human EST encoded protein SEQ ID NO: 1847.
 XX
 KW Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
 KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
 KW diagnostics; forensic test; gene mapping; genetic disorder; biodiversity;
 KW gene therapy; nutrition.
 XX
 OS Homo sapiens.
 XX
 PN WO200154477-A2.
 XX
 PD 02-AUG-2001.
 XX
 PF 25-JAN-2001; 2001WO-US002687.
 XX
 PR 25-JAN-2000; 2000US-00491404.
 PR 17-JUL-2000; 2000US-00617746.
 PR 03-AUG-2000; 2000US-00631451.
 PR 15-SEP-2000; 2000US-00663870.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
 PI Cao Y, Drmanac RA, Zhang J, Werhman T;
 DR WPI; 2001-476164/51.
 DR N-PSDB; AAH98981.
 XX
 PT Isolated polypeptide for treatment of diseases, diagnostics, raising
 PT antibodies and research use.
 XX
 PS Claim 20; Page 1198-1201; 1275pp; English.
 XX
 CC The present invention provides the protein and coding sequences of novel
 CC proteins from a variety of organisms, including human, dog, cat, horse,
 CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
 CC urchin and tomato. These were derived from expressed sequence tags (ESTs)
 CC from the organism of interest. They can be used in diagnostics,
 CC forensics, gene mapping, identification of mutations, to assess
 CC biodiversity and for nutritional purposes. The present sequence is a
 CC protein of the invention

XX
 SQ Sequence 1299 AA;

Query Match 86.8%; Score 33; DB 4; Length 1299;
 Best Local Similarity 85.7%; Pred. No. 1.4e+03;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSPS 7
 Db 1171 WMLSPS 1177

RESULT 21
 ID ADK67912 standard; protein; 1311 AA.
 AC ADK67912;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE Human extracellular messenger (EXMES) polypeptide.
 XX
 KW Human; extracellular messenger; EXMES; respiratory-gen.; anti-allergic;
 KW asthmatic; anti-inflammatory; antidiabetic; neuroprotective;
 KW muscular-gen.; anti-arthritic; osteopathic; hepatotropic; antiparasitic;
 KW virucide; fungicide; antiparasitic; protozoacide; antihelminthic;
 KW cyostatic; gene therapy.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..30
 FT /label=Signal_peptide
 FT /note="Spans residues 1 to 18, 20, 21, 24, 29 or 30
 FT according to identification method"
 PN WO2004013292-A2.
 XX
 PD 12-FEB-2004.
 XX
 PF 30-JUL-2003; 2003WO-US024084.
 XX
 PR 02-AUG-2002; 2002US-0400810P.
 PR 19-SEP-2002; 2002US-0412197P.
 PR 04-OCT-2002; 2002US-0416004P.
 PR 08-NOV-2002; 2002US-0424862P.
 XX
 PA (INCY-) INCYTE CORP.
 XX
 PI Elliott VS, Khare R, Tran UK, Swarnakar A, Marquis JP;
 PI Richardson TW, Emerling BM, Lindquist EA, Chawla NK, Ramkumar J;
 PI Lee SY;
 DR WPI; 2004-157116/15.
 DR N-PSDB; ADK67917.
 XX
 PT New extracellular messengers and nucleic acids, useful for diagnosing,
 PT treating or preventing e.g. allergies, anemia, Crohn's disease, diabetes
 PT mellitus, myasthenia gravis, osteoporosis, psoriasis, cancers or
 PT autoimmune thyroiditis.
 XX
 PS Claim 60; SEQ ID NO 5; 165pp; English.
 XX
 CC The present sequence is that of novel human extracellular messenger
 CC (EXMES) Incyte ID NO: 7513018CD1 polypeptide. The protein shows homology
 CC to human megakaryocyte stimulating factor. The invention provides EXMES
 CC polynucleotides and polypeptides, as well as expression vectors, host
 CC cells, antibodies, agonists and antagonists, and methods for diagnosing,
 CC treating or preventing disorders associated with aberrant expression of
 CC EXMES, especially autoimmune and inflammatory disorders, cell
 CC proliferative disorders and endocrine disorders, e.g. adult respiratory
 CC distress syndrome, allergy, anaemia, erythroblastosis foetalis, Crohn's
 CC disease, diabetes mellitus, myasthenia gravis, osteoarthritis,

CC osteoporosis, pancreatitis, psoriasis, viral, bacterial, fungal,
CC parasitic, protozoal or helminthic infections, cancers, autoimmune
CC thyroiditis, cretinism, Plummer's disease or thyroid carcinoma.
CC Embodiments also provide methods for using the purified EXMSS and/or
CC their encoding polynucleotides for facilitating the drug discovery
CC process, including determining of efficacy, dosage, toxicity and
CC pharmacology, and for investigating the pathogenesis of diseases and
CC medical conditions.
XX
SQ Sequence 1311 AA;

Query Match 86.8%; Score 33; DB 8; Length 1311;
Best Local Similarity 85.7%; Pred. No. 1.4e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 WMLSPFS 7
Db 1078 WMLSPFS 1084

RESULT 22

ADK65819
ID ADK65819 standard; protein; 1320 AA.

AC ADK65819;

DT 06-MAY-2004 (first entry)

DE Angiogenesis-differentially expressed protein ANH0316.

KM cytostatic; cardiast; vasotropic; antiarteriosclerotic;
KM angiogenesis inhibitor; angiogenesis stimulator; angiogenic index;
KM gene expression; cancer; coronary artery disease; myocardial ischemia;
KM coronary arteriosclerosis; forensic medicine.
XX

OS Homo sapiens.

PN WO2003066831-A2.

PD 14-AUG-2003.

PF 07-FEB-2003; 2003WO-US003848.

PR 07-FEB-2002; 2002US-00067482.

PR 10-JUN-2002; 2002US-00164595.

PR 16-AUG-2002; 2002US-0403649P.

PR 03-JAN-2003; 2003US-0437746P.

XX (ORIG-) ORIGENE TECHNOLOGIES INC.

PI Sun Z, Li X, Kovacs KF, Fan W, Jay G;

DR WPI; 2003-731502/69.

DR N-PSDB; ADK65818.

PT Determining the angiogenic index of a tissue or cell sample using
XX expression levels of differentially expressed genes, useful for
XX diagnosing or treating cancer, coronary artery disease, myocardial
XX ischemia and/or arteriosclerosis.

PS Claim 23; SEQ ID NO 58; 296bp; English.

XX The invention relates to a method of determining the angiogenic index of
CC a tissue or cell sample comprising assessing, in a sample, the expression
CC levels of one or more differentially-expressed gene from any of 34 DNA
CC sequences, given in the specification, where the levels are indicative of
CC the angiogenic index. The methods and compositions of the present
CC invention are useful for diagnosing, preventing and/or treating cancer,
CC coronary artery disease, myocardial ischemia or coronary
CC arteriosclerosis. They can also be used in research, drug discovery and
CC forensic medicine involving angiogenesis. This sequence corresponds to
CC one of the differentially expressed proteins of the invention.
XX

SQ Sequence 1320 AA;

Query Match 86.8%; Score 33; DB 7; Length 1320;
Best Local Similarity 85.7%; Pred. No. 1.4e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 WMLSPFS 7
Db 1087 WMLSPFS 1093

RESULT 23

AAR26049
ID AAR26049 standard; protein; 1404 AA.

AC AAR26049;

DT 25-MAR-2003 (revised)

DT 02-FEB-1993 (first entry)

DE MSF precursor.

KM Megakaryocyte colony stimulating factor; secretion signal; meg-CSF;
KM stability; proteolytic cleavage; adhesion; alternative splicing.
XX

OS Synthetic.

PH Key Location/Qualifiers

FT Region 1..26

FT Region /label= Exon_I

FT Region /label= Exon_II

FT Region /label= Exon_III

FT Region /label= Exon_IV

FT Region /label= Exon_V

FT Region /label= Exon_VI

FT Region /label= Exon_VII

FT Region /label= Exon_VIII

FT Region /label= Exon_IX

FT Region /label= Exon_X

FT Region /label= Exon_XI

FT Region /label= Exon_XII

FT Region /label= Exon_XIII

FT Region /label= Exon_XIV

FT Region /label= Exon_XV

FT Region /label= Exon_XVI

FT Region /label= Exon_XVII

FT Region /label= Exon_XVIII

FT Region /label= Exon_XIX

FT Region /label= Exon_XX

FT Region /label= Exon_XXI

FT Region /label= Exon_XXII

FT Region /label= Exon_XXIII

FT Region /label= Exon_XXIV

FT Region /label= Exon_XXV

FT Region /label= Exon_XXVI

FT Region /label= Exon_XXVII

FT Region /label= Exon_XXVIII

FT Region /label= Exon_XXIX

FT Region /label= Exon_XXX

FT Region /label= Exon_XXXI

FT Region /label= Exon_XXXII

FT Region /label= Exon_XXXIII

FT Region /label= Exon_XXXIV

FT Region /label= Exon_XXXV

FT Region /label= Exon_XXXVI

FT Region /label= Exon_XXXVII

FT Region /label= Exon_XXXVIII

FT Region /label= Exon_XXXIX

FT Region /label= Exon_XXXX

FT Region /label= Exon_XXXXI

FT Region /label= Exon_XXXXII

FT Region /label= Exon_XXXXIII

FT Region /label= Exon_XXXXIV

FT Region /label= Exon_XXXXV

FT Region /label= Exon_XXXXVI

FT Region /label= Exon_XXXXVII

FT Region /label= Exon_XXXXVIII

FT Region /label= Exon_XXXXIX

FT Region /label= Exon_XXXXX

FT Region /label= Exon_XXXXXI

FT Region /label= Exon_XXXXXII

FT Region /label= Exon_XXXXXIII

FT Region /label= Exon_XXXXXIV

FT Region /label= Exon_XXXXXV

FT Region /label= Exon_XXXXXVI

FT Region /label= Exon_XXXXXVII

FT Region /label= Exon_XXXXXVIII

FT Region /label= Exon_XXXXXIX

FT Region /label= Exon_XXXXXX

FT Region /label= Exon_XXXXXXI

FT Region /label= Exon_XXXXXXII

FT Region /label= Exon_XXXXXXIII

FT Region /label= Exon_XXXXXXIV

FT Region /label= Exon_XXXXXXV

FT Region /label= Exon_XXXXXXVI

FT Region /label= Exon_XXXXXXVII

FT Region /label= Exon_XXXXXXVIII

FT Region /label= Exon_XXXXXXIX

FT Region /label= Exon_XXXXXXX

FT Region /label= Exon_XXXXXXXI

FT Region /label= Exon_XXXXXXXII

FT Region /label= Exon_XXXXXXXIII

FT Region /label= Exon_XXXXXXXIV

FT Region /label= Exon_XXXXXXXV

FT Region /label= Exon_XXXXXXXVI

FT Region /label= Exon_XXXXXXXVII

FT Region /label= Exon_XXXXXXXVIII

FT Region /label= Exon_XXXXXXXIX

FT Region /label= Exon_XXXXXXX

FT Region /label= Exon_XXXXXXXI

FT Region /label= Exon_XXXXXXXII

FT Region /label= Exon_XXXXXXXIII

FT Region /label= Exon_XXXXXXXIV

FT Region /label= Exon_XXXXXXXV

FT Region /label= Exon_XXXXXXXVI

FT Region /label= Exon_XXXXXXXVII

FT Region /label= Exon_XXXXXXXVIII

FT Region /label= Exon_XXXXXXXIX

FT Region /label= Exon_XXXXXXX

FT Region /label= Exon_XXXXXXXI

FT Region /label= Exon_XXXXXXXII

FT Region /label= Exon_XXXXXXXIII

FT Region /label= Exon_XXXXXXXIV

FT Region /label= Exon_XXXXXXXV

FT Region /label= Exon_XXXXXXXVI

FT Region /label= Exon_XXXXXXXVII

FT Region /label= Exon_XXXXXXXVIII

FT Region /label= Exon_XXXXXXXIX

FT Region /label= Exon_XXXXXXX

FT Region /label= Exon_XXXXXXXI

FT Region /label= Exon_XXXXXXXII

FT Region /label= Exon_XXXXXXXIII

FT Region /label= Exon_XXXXXXXIV

FT Region /label= Exon_XXXXXXXV

FT Region /label= Exon_XXXXXXXVI

FT Region /label= Exon_XXXXXXXVII

FT Region /label= Exon_XXXXXXXVIII

FT Region /label= Exon_XXXXXXXIX

FT Region /label= Exon_XXXXXXX

FT Region /label= Exon_XXXXXXXI

FT Region /label= Exon_XXXXXXXII

FT Region /label= Exon_XXXXXXXIII

FT Region /label= Exon_XXXXXXXIV

FT Region /label= Exon_XXXXXXXV

FT Region /label= Exon_XXXXXXXVI

FT Region /label= Exon_XXXXXXXVII

FT Region /label= Exon_XXXXXXXVIII

FT Region /label= Exon_XXXXXXXIX

FT Region /label= Exon_XXXXXXX

FT Region /label= Exon_XXXXXXXI

FT Region /label= Exon_XXXXXXXII

FT Region /label= Exon_XXXXXXXIII

FT Region /label= Exon_XXXXXXXIV

FT Region /label= Exon_XXXXXXXV

FT Region /label= Exon_XXXXXXXVI

FT Region /label= Exon_XXXXXXXVII

FT Region /label= Exon_XXXXXXXVIII

FT Region /label= Exon_XXXXXXXIX

FT Region /label= Exon_XXXXXXX

FT Region /label= Exon_XXXXXXXI

FT Region /label= Exon_XXXXXXXII

FT Region /label= Exon_XXXXXXXIII

FT Region /label= Exon_XXXXXXXIV

FT Region /label= Exon_XXXXXXXV

FT Region /label= Exon_XXXXXXXVI

FT Region /label= Exon_XXXXXXXVII

FT Region /label= Exon_XXXXXXXVIII

FT Region /label= Exon_XXXXXXXIX

FT Region /label= Exon_XXXXXXX

FT Region /label= Exon_XXXXXXXI

FT Region /label= Exon_XXXXXXXII

FT Region /label= Exon_XXXXXXXIII

FT Region /label= Exon_XXXXXXXIV

FT Region /label= Exon_XXXXXXXV

FT Region /label= Exon_XXXXXXXVI

FT Region /label= Exon_XXXXXXXVII

FT Region /label= Exon_XXXXXXXVIII

FT Region /label= Exon_XXXXXXXIX

FT Region /label= Exon_XXXXXXX

FT Region /label= Exon_XXXXXXXI

FT Region /label= Exon_XXXXXXXII

FT Region /label= Exon_XXXXXXXIII

FT Region /label= Exon_XXXXXXXIV

FT Region /label= Exon_XXXXXXXV

FT Region /label= Exon_XXXXXXXVI

FT Region /label= Exon_XXXXXXXVII

FT Region /label= Exon_XXXXXXXVIII

FT Region /label= Exon_XXXXXXXIX

FT Region /label= Exon_XXXXXXX

FT Region /label= Exon_XXXXXXXI

FT Region /label= Exon_XXXXXXXII

FT Region /label= Exon_XXXXXXXIII

FT Region /label= Exon_XXXXXXXIV

FT Region /label= Exon_XXXXXXXV

FT Region /label= Exon_XXXXXXXVI

FT Region /label= Exon_XXXXXXXVII

FT Region /label= Exon_XXXXXXXVIII

FT Region /label= Exon_XXXXXXXIX

FT Region /label= Exon_XXXXXXX

FT Region /label= Exon_XXXXXXXI

FT Region /label= Exon_XXXXXXXII

FT Region /label= Exon_XXXXXXXIII

FT Region /label= Exon_XXXXXXXIV

FT Region /label= Exon_XXXXXXXV

FT Region /label= Exon_XXXXXXXVI

FT Region /label= Exon_XXXXXXXVII

FT Region /label= Exon_XXXXXXXVIII

FT Region /label= Exon_XXXXXXXIX

FT Region /label= Exon_XXXXXXX

FT Region /label= Exon_XXXXXXXI

FT Region /label= Exon_XXXXXXXII

FT Region /label= Exon_XXXXXXXIII

FT Region /label= Exon_XXXXXXXIV

FT Region /label= Exon_XXXXXXXV

FT Region /label= Exon_XXXXXXXVI

FT Region /label= Exon_XXXXXXXVII

FT Region /label= Exon_XXXXXXXVIII

FT Region /label= Exon_XXXXXXXIX

FT Region /label= Exon_XXXXXXX

FT Region /label= Exon_XXXXXXXI

FT Region /label= Exon_XXXXXXXII

FT Region /label= Exon_XXXXXXXIII

FT Region /label= Exon_XXXXXXXIV

FT Region /label= Exon_XXXXXXXV

FT Region /label= Exon_XXXXXXXVI

FT Region /label= Exon_XXXXXXXVII

FT Region /label= Exon_XXXXXXXVIII

FT Region /label= Exon_XXXXXXXIX

FT Region /label= Exon_XXXXXXX

FT Region /label= Exon_XXXXXXXI

FT Region /label= Exon_XXXXXXXII

FT Region /label= Exon_XXXXXXXIII

FT Region /label= Exon_XXXXXXXIV

FT Region /label= Exon_XXXXXXXV

FT Region /label= Exon_XXXXXXXVI

FT Region /label= Exon_XXXXXXXVII

FT Region /label= Exon_XXXXXXXVIII

FT Region /label= Exon_XXXXXXXIX

FT Region /label= Exon_XXXXXXX

FT Region /label= Exon_XXXXXXXI

FT Region /label= Exon_XXXXXXXII

FT Region /label= Exon_XXXXXXXIII

FT Region /label= Exon_XXXXXXXIV

FT Region /label= Exon_XXXXXXXV

FT Region /label= Exon_XXXXXXXVI

FT Region /label= Exon_XXXXXXXVII

FT Region /label= Exon_XXXXXXXVIII

FT Region /label= Exon_XXXXXXXIX

FT Region /label= Exon_XXXXXXX

FT Region /label= Exon_XXXXXXXI

FT Region /label= Exon_XXXXXXXII

FT Region /label= Exon_XXXXXXXIII

FT Region /label= Exon_XXXXXXXIV

FT Region /label= Exon_XXXXXXXV

FT Region /label= Exon_XXXXXXXVI

FT Region /label= Exon_XXXXXXXVII

FT Region /label= Exon_XXXXXXXVIII

FT Region /label= Exon_XXXXXXXIX

FT Region /label= Exon_XXXXXXX

FT Region /label= Exon_XXXXXXXI

FT Region /label= Exon_XXXXXXXII

FT Region /label= Exon_XXXXXXXIII

FT Region /label= Exon_XXXXXXXIV

FT Region /label= Exon_XXXXXXXV

FT Region /label= Exon_XXXXXXXVI

FT Region /label= Exon_XXXXXXXVII

FT Region /label= Exon_XXXXXXXVIII

FT Region /label= Exon_XXXXXXXIX

FT Region /label= Exon_XXXXXXX

FT Region /label= Exon_XXXXXXXI

FT Region /label= Exon_XXXXXXXII

FT Region /label= Exon_XXXXXXXIII

FT Region /label= Exon_XXXXXXXIV

FT Region /label= Exon_XXXXXXXV

PS Claim 1, 2 and 3; Fig 1; 87pp; English.
XX
CC The sequence given is a full length translation from the megakaryocyte
CC stimulating factor (MSF) precursor. The sequence covered by exons II, III
CC and IV encodes megakaryocyte stimulating factor (MSF). This sequence is
CC modified by the addition of an N-terminal sequence encoding a secretory
CC leader, an initiating methionine preceding exon II and a terminating
CC codon following exon IV. The cDNA sequence given contains sequences
CC derived from human megakaryocyte colony stimulating factor (meg-CSF).
CC Exon I contains the initiating methionine, and encodes a classical
CC mammalian protein secretion signal sequence. The sequence encoding the
CC original meg-CSF includes exons II-IV and is thought to terminate in the
CC region between amino acid residues 134 - 147. The primary transcript of
CC this gene may be cleaved in different ways to yield a family of mRNAs
CC each encoding a different MSF protein. Exons V and VI are thought to be
CC related to the activity of the factor and are also implicated in the
CC stability, folding and processing of the molecule. These exons are also
CC thought to play a role in the observed synergy of MSF with other
CC cytokines. Exons V - XII are believed to be implicated in the processing
CC or folding of the appropriate structure of the resulting factor, ie. one
CC or more of these exons may contain sequences which direct proteolytic
CC cleavage, adhesion, organisation of the cellular matrix or extracellular
CC matrix processing. Both naturally occurring and non-naturally occurring
CC MSF's may be characterised by various combinations of alternatively
CC spliced exons from this sequence, with the exons spliced together in
CC differing orders to form different members of the MSF family. (Updated on
CC 25-MAR-2003 to correct PN field.)
CC
XX
SQ Sequence 1404 AA;

Query Match 86.8%; Score 33; DB 2; Length 1404;
Best Local Similarity 85.7%; Pred. No. 1.5e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
Db 1171 WMLSPFS 1177

RESULT 24
AAB29773
ID AAB29773 standard; protein; 1404 AA.
XX
AC AAB29773;
XX
DT 28-FEB-2001 (first entry)
XX
DE Human megakaryocyte stimulating factor (MSF), SEQ ID NO:1.
XX
KW Human MSF; megakaryocyte stimulating factor; tribonectin;
KW alternative splicing; joint boundary lubricant; O-linked oligosaccharide;
KW osteoarthritis; tribosupplementation; tissue adhesion inhibition;
KW friction coefficient reduction; gene therapy; antiarthritic; osteopathic.
XX
OS Homo sapiens.
XX
PN WO200064930-A2.
XX
PD 02-NOV-2000.
XX
PF 24-APR-2000; 2000WO-US010953.
XX
PR 23-APR-1999; 99US-00298970.
XX
PA (RHOD-) RHODE ISLAND HOSPITAL LIFESPAN PARTNER.
XX
PI Jay GD;
XX
DR WPI, 2001-024673/03.
XX
DR N-PSDB; AAC81498.
XX
PT Novel tribonectin polypeptide useful as lubricant for treating
PT osteoarthritis, comprises O-linked lubricating moiety.

XX
PS Claim 3; Page 7; 47pp; English.
XX
CC The invention relates to a human tribonectin which is a product of
CC alternative splicing of the human MSF (megakaryocyte stimulating factor)
CC gene. The tribonectin has at least one O-linked oligosaccharide
CC lubricating moiety and has a polypeptide sequence comprising 1-76 repeats
CC of a motif having at least 50% identity to the sequence KEDAPPT
CC (AAB29774). The invention also relates to a nucleic acid encoding a human
CC MSF-derived tribonectin; a biocompatible composition comprising a human
CC tribonectin for inhibiting tissue adhesion formation; and a method of
CC diagnosing osteoarthritis or a predisposition to osteoarthritis by
CC measuring the amount of MSF or its fragment in a biological sample of a
CC mammal, wherein an increased amount of MSF compared to a control
CC indicates the presence of or predisposition to developing osteoarthritis.
CC The tribonectin and DNA encoding it are useful in the treatment of
CC osteoarthritis, where they may be used for lubricating mammalian joints,
CC such as articulating joints of humans, dogs or horses. The tribonectin,
CC when formulated as a membrane, foam, gel or fibre, is useful for
CC inhibiting adhesion between two surfaces such as the injured tissues of a
CC mammal, where the injury is caused by a surgical insertion or trauma, or
CC an artificial device e.g., an orthopaedic implant. In particular, one of
CC the surfaces is pericardial tissue. DNA encoding a tribonectin may be
CC used in gene therapy. The present sequence represents human MSF
XX
SQ Sequence 1404 AA;

Query Match 86.8%; Score 33; DB 4; Length 1404;
Best Local Similarity 85.7%; Pred. No. 1.5e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
Db 1171 WMLSPFS 1177

RESULT 25
AAB60568
ID AAB60568 standard; protein; 1404 AA.
XX
AC AAB60568;
XX
DT 27-APR-2001 (first entry)
XX
DE Human megakaryocyte stimulating factor (MSF, CACP).
XX
KW Human; CACP protein; campodactylly-arthropathy-coxa vara-pericarditis;
KW MSF; megakaryocyte stimulating factor; synovial lubricant;
KW chromosome 1q25-31; osteoarthritis; joint lubrication; osteopathic;
KW antiarthritic.
XX
OS Homo sapiens.
XX
PN WO200107068-A1.
XX
PD 01-FEB-2001.
XX
PF 21-JUL-2000; 2000WO-US020002.
XX
PR 23-JUL-1999; 99US-0145328P.
XX
PR 19-JUL-2000; 2000US-00145326.
XX
PA (UYCA-) UNIV CASE WESTERN RESERVE.
XX
PI Warman ML;
XX
DR WPI, 2001-182721/18.
XX
PT New composition comprising the campodactylly-arthropathy-coxa vara-
PT pericarditis protein in combination with an anesthetic, useful for
PT treating osteoarthritis, or as lubricants of tissue and joints.
XX
PS Example 1; Page; 34pp; English.

QY 1 WMLSAFS 7
 |||||
 DB 1171 WMLSPFS 1177

RESULT 28
 ADX69335

ID ADX69335 standard; protein; 1404 AA.

AC ADX69335;

DE 05-MAY-2005 (first entry)

XX Human heparin binding protein #35.

XX Heparin binding protein; Neuroprotective; Nootropic; Antiparkinsonian;
 XX Anticonvulsant; VEGF-3 receptor; Angiogenesis stimulator; Gene Therapy;
 XX Vascular endothelial growth factor receptor 3; VEGFR-3;
 XX angiogenesis disorder; neurodegenerative disorder; Alzheimers disease;
 XX Parkinsons disease; motor neurone disease; dementia; paralysis; VEGF-C;
 XX neurological disease; Huntingtons chorea;
 XX vascular endothelial growth factor receptor 3; neurological disease.

OS Homo sapiens.

PN WO2005016963-A2.

PD 24-FEB-2005.

PF 14-JUN-2004; 2004WO-US019122.

PR 12-JUN-2003; 2003US-0478114P.

PR 12-JUN-2003; 2003US-0478390P.

PR 23-SEP-2003; 2003US-00669176.

PA (LUDW-) LUDWIG INST CANCER RES.
 (LICN) LICENTIA LTD.

PI Alitalo K, He Y, Tammela T;

DR WPI; 2005-182331/19.

XX New heparin-binding Vascular Endothelial Growth Factor Receptor 3 ligands
 PT comprising prepro-VEGF-C, prepro-VEGF-D sequences or fragments, for
 PT treating neurodegenerative disorder, e.g. Alzheimer's, Parkinson's, or
 PT Huntington's disease.

PS Disclosure; SEQ ID NO 55; 219pp; English.

XX The invention relates to heparin-binding vascular endothelial growth
 CC factor receptor 3 (VEGFR-3) proteins and encoding polynucleotides. The
 CC heparin binding VEGFR-3 proteins are used for stimulating
 CC lymphangiogenesis or angiogenesis in a mammal, and for modulating the
 CC growth of mammalian endothelial cells, mammalian endothelial precursor
 CC cells or hematopoietic progenitor cells. The polypeptide may also be used
 CC for promoting recruitment, proliferation, differentiation, migration or
 CC survival of neuronal cells or neuronal precursor cells, and for treating
 CC neurodegenerative disorder, e.g. Alzheimer's disease, Parkinson's
 CC disease, Huntington's disease, motor neuron disease, Amyotrophic Lateral
 CC Sclerosis (ALS), dementia, or cerebral palsy. The present sequence
 CC represents a human heparin binding protein of the invention.

XX Sequence 1404 AA;

QY Query Match 86.8%; Score 33; DB 9; Length 1404;

Best Local Similarity 85.7%; Pred. No. 1.5e+03;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
 |||||
 DB 1171 WMLSPFS 1177

RESULT 29
 ADY15658
 ID ADY15658 standard; protein; 1404 AA.

AC ADY15658;

DE 05-MAY-2005 (first entry)

XX PRO polypeptide SEQ ID NO 1464.

XX Antiinflammatory; Immune disorder; Dermatological; Immunosuppressive;
 XX Antirheumatic; Antiarthritic; Osteopathic; Hemostatic; Antianemic;
 XX Antithyroid; Antidiabetic; Nephrotropic; CNS-Gen.; Hepatotropic;
 XX Vnucide; Gastrointestinal-Gen.; Antiporiatic; Antiasthmatic;
 XX Antiallergic; ds; gene; diagnosis.

OS Homo sapiens.

PN WO2005016962-A2.

PD 24-FEB-2005.

PF 11-AUG-2004; 2004WO-US026249.

PR 11-AUG-2003; 2003US-0493546P.

PA (GETH) GENENTECH INC.

PI Abbas A, Clark H, Ouyang W, Williams MP, Wood WT, Wu TD;

DR WPI; 2005-182330/19.

XX New nucleic acid encoding PRO polypeptide, useful for diagnosing and
 PT treating an immune related disorder, e.g. systemic lupus erythematosus,
 PT rheumatoid arthritis, osteoarthritis, thyroiditis, or diabetes mellitus.
 PS Claim 8; SEQ ID NO 1464; 158pp; English.

XX The invention relates to an isolated nucleic acid encoding a PRO
 CC polypeptide. The polypeptide, agonist or an antagonist, antibody,
 CC composition, and method are useful for diagnosing and treating an immune
 CC related disorder, e.g. systemic lupus erythematosus, rheumatoid
 CC arthritis. The present sequence represents a DNA encoding a PRO
 CC polypeptide.

XX Sequence 1404 AA;

QY Query Match 86.8%; Score 33; DB 9; Length 1404;

Best Local Similarity 85.7%; Pred. No. 1.5e+03;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
 |||||
 DB 1171 WMLSPFS 1177

RESULT 30

ADX66477
 ID AED66477 standard; protein; 1404 AA.

AC AED66477;

DE 29-DEC-2005 (first entry)

XX Human lubricin protein sequence.

XX Wound healing agent; tissue regeneration; antisense therapy; trauma;
 XX vulnereary; injury; fractures; osteopathic; synovitis; inflammation;
 XX lubricin.

OS Homo sapiens.

PN WO2005102363-A2.

[illegible]

```

XX PT Jay G;
XX XX
XX DR WPI; 2006-136979/14.
XX DR N-PSDB; AEF54257.
XX XX
XX PT Visco-supplementation composition, useful for the lubrication and
XX PT chondroprotection of mammalian joint, comprising hyaluronic acid and
XX PT tribonectin.
XX PS
XX PS Disclosure; SEQ ID NO 1; 50pp; English.
XX XX
XX CC The invention relates to a viscosupplementation composition (A) which
XX CC comprises hyaluronic acid (1-5 mg/ml) and tribonectin (10-250 mu g/ml).
XX CC Also described is the use of tribonectin in the preparation of a
XX CC medicament for the lubrication and chondroprotection of a mammalian
XX CC joint, where the tribonectin is added to a visco-supplement in the
XX CC preparation of the medicament, which increases the elasticity of the
XX CC visco-supplement. (A) is useful in the preparation of medicament for the
XX CC lubrication or chondroprotection of a mammalian joint (particularly
XX CC articulating a joint of a human, horse or dog). The present sequence
XX CC represents human megakaryocyte stimulating factor (MSF).
XX SQ
XX SQ Sequence 1404 AA;

Query Match 86.8%; Score 33; DB 10; Length 1404;
Best Local Similarity 85.7%; Pred. No. 1.5e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 WMLSAFS 7
Db 1171 WMLSPFS 1177

RESULT 32
AEF89839
ID AEF89839 standard; protein; 1404 AA.
XX AEF89839;
XX AC
XX DT 20-APR-2006 (first entry)
XX XX
XX DE Human megakaryocyte stimulating factor (MSF) - SEQ ID 1.
XX XX
XX KW tribonectin; arthritis; antiarthritic; osteoarthritis; osteopathic;
XX KW megakaryocyte stimulating factor; MSF.
XX OS
XX OS Homo sapiens.
XX PN US7001881-B1.
XX PD
XX PD 21-FEB-2006.
XX PF
XX PF 24-APR-2000; 2000US-00556246.
XX PR
XX PR 23-APR-1999; 99US-00298970.
XX PA (RHOD-) RHODE ISLAND HOSPITAL.
XX PI
XX PI Jay GD;
XX DR WPI; 2006-170699/18.
XX DR N-PSDB; AEF89840.
XX XX
XX PT New isolated tribonectin comprising a boundary-lubricating amount of
XX PT polypeptide, and an O-linked oligosaccharide group, for use in treating
XX PT arthritic diseases, and to coat artificial limbs and joints.
XX PS
XX PS Claim 1; SEQ ID NO 1; 34pp; English.
XX XX
XX CC The invention comprises a tribonectin protein (a boundary lubricating
XX CC polypeptide) the tribonectin protein of the invention comprises the
XX CC human megakaryocyte stimulating factor (MSF) protein. The tribonectin

```

CC protein of the invention is useful for treating arthritic disease and
CC osteoarthritis disease, and for coating artificial limbs and joints prior
CC to implantation. The present amino acid sequence represents the human MSF
CC protein of the invention.
XX

SQ Sequence 1404 AA;

Query Match 86.8%; Score 33; DB 10; Length 1404;

Best Local Similarity 85.7%; Pred. No. 1.5e+03;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7

Db 1171 WMLSPFS 1177

RESULT 33

AAU32262

XX AAU32262 standard; protein; 1415 AA.

AC AAU32262;

XX 18-DEC-2001 (first entry)

DT 18-DEC-2001 (first entry)

XX Novel human secreted protein #2753.

DE Human; vaccination; gene therapy; nutritional supplement;

XX Human; vaccination; gene therapy; nutritional supplement;

KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;

XX immune suppression; immune stimulation; anti-inflammatory; leukaemia.

OS Homo sapiens.

XX Homo sapiens.

PN WO200179449-A2.

XX 25-OCT-2001.

PD 16-APR-2001; 2001WO-US008656.

XX 18-APR-2000; 2000US-00552929.

PR 26-JAN-2001; 2001US-00770160.

XX (HYSE-) HYSEQ INC.

PA Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-611725/70.

DR Nucleic acids encoding a range of human polypeptides, useful in genetic

XX vaccination, testing and therapy.

PT Claim 20; Page 573; 765pp; English.

XX The invention relates to novel human secreted polypeptides. The

XX polypeptides and antibodies to the polypeptides are useful for

CC determining the presence of or predisposition to a disease associated

CC with altered levels of polypeptide. The polypeptides are also useful for

CC identifying agents (agonists and antagonists) that bind to them. Cells

CC expressing the proteins are useful for identifying a therapeutic agent

CC for use in treatment of a pathology related to aberrant expression or

CC physiological interactions of the polypeptide. Vectors comprising the

CC nucleic acids encoding the polypeptides and cells genetically engineered

CC to express them are also useful for producing the proteins. The proteins

CC are useful in genetic vaccination, testing and therapy, and can be used

CC as nutritional supplements. They may be used to increase stem cell

CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon

CC and/or nerve tissue growth or regeneration, immune suppression and/or

CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.

CC AAU29510-AAU33304 represents the amino acid sequences of novel human

CC secreted proteins of the invention

XX Sequence 1415 AA;

SQ Query Match 86.8%; Score 33; DB 4; Length 1415;

Best Local Similarity 85.7%; Pred. No. 1.5e+03;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7

Db 1182 WMLSPFS 1188

RESULT 34

AAM82730

XX AAM82730 standard; protein; 31 AA.

AC AAM82730;

XX 07-NOV-2001 (first entry)

DT 07-NOV-2001 (first entry)

XX Human immune/haematopoietic antigen SEQ ID NO:10323.

DE Human immune/haematopoietic antigen SEQ ID NO:10323.

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;

KW cytostatic; gene therapy; vaccine; metastasis.

XX Homo sapiens.

OS Homo sapiens.

PN WO200157182-A2.

XX 09-AUG-2001.

PD 17-JAN-2001; 2001WO-US001354.

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

XX 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

XX 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

XX 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

XX 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

XX 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

XX 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

XX 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

XX 26-JUL-2000; 2000US-0220963P.

PR 14-AUG-2000; 2000US-0224518P.

XX 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

XX 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 14-AUG-2000; 2000US-0225759P.

PR 18-AUG-2000; 2000US-0226279P.

PR 22-AUG-2000; 2000US-0226681P.

PR 22-AUG-2000; 2000US-0226686P.

PR 22-AUG-2000; 2000US-0227182P.

PR 23-AUG-2000; 2000US-0227009P.

PR 30-AUG-2000; 2000US-0228924P.

PR 01-SEP-2000; 2000US-0229287P.

PR 01-SEP-2000; 2000US-0229343P.

PR 01-SEP-2000; 2000US-0229344P.

PR 01-SEP-2000; 2000US-0229345P.

PR 05-SEP-2000; 2000US-0229509P.

PR 05-SEP-2000; 2000US-0229513P.

PR 06-SEP-2000; 2000US-0230437P.

PR 06-SEP-2000; 2000US-0230438P.

PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.

PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249246P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 DR WPI; 2001-483426/52.
 DR N-PSDB; AAK55511.
 XX
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and metastasis.
 XX
 PS Claim 11; SEQ ID NO 10323; 3071pp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytosolic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patients own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting the
 CC nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AAK54950 and AAM82169
 CC represent sequences used in the exemplification of the present invention
 CC
 XX
 SQ Sequence 31 AA;
 Query Match 84.2%; Score 32; DB 4; Length 31;
 Best Local Similarity 85.7%; Pred. No. 39;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 WMLSAFS 7
 Db 6 WMLSAFS 12
 RESULT 35
 ABP04103
 ID ABP04103 standard; protein; 81 AA.
 XX
 AC ABP04103;
 XX
 DT 25-JUN-2002 (first entry)
 XX
 DE Human ORFX protein sequence SEQ ID NO:8188.
 XX
 KW Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
 KW hyperproliferative disorder; psoriasis; benign tumor; haemorrhage;

KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
 KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
 KW hypertension; hypothyroidism; cholesterol ester storage disease;
 KW immune deficiency; immune disorder; infectious disease;
 KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
 KW myasthenia gravis.
 XX
 OS Homo sapiens.
 PN WO200192523-A2.
 PD 06-DEC-2001.
 XX
 PF 29-MAY-2001; 2001WO-US010836.
 XX
 PR 30-MAY-2000; 2000US-0206132P.
 PR 29-AUG-2000; 2000US-0228716P.
 XX
 PA (CURA-) CURAGEN CORP.
 PI Shimkets RA, Leach MD;
 XX
 DR WPI; 2002-106308/14.
 DR N-PSDB; ABN19855.
 XX
 PT Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and autoimmune disorders.
 XX
 PS Disclosure; SEQ ID NO 8188; 1037bp; English.
 XX
 CC The present invention describes substantially purified human proteins
 CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
 CC in the specification). ABN15762 to ABN27252 encode the human ORFX
 CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
 CC treating or preventing a pathology associated with an ORFX-associated
 CC disorder in humans, and in the manufacture of a medicament for treating a
 CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
 CC sequences can be used in gene therapy. ORFX sequences can be used in the
 CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
 CC psoriasis, benign tumors, keloid, degenerative disorders, haemorrhage,
 CC osteoarthritis, neurodegenerative disorders, disorders related to organ
 CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
 CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
 CC storage disease, various immune deficiencies and disorders, infectious
 CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
 CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
 CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
 CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
 CC bone degenerative disorders, or periodontal disease, and for gut
 CC protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues and conditions resulting from
 CC systemic cytokine damage. N.B. The sequence data for this patent did not
 CC form part of the printed specification, but was obtained in electronic
 CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 81 AA;
 QY
 Db Query Match 84.2%; Score 32; DB 5; Length 81;
 Best Local Similarity 71.4%; Pred. No. 1.1e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 1 WMLSAFS 7
 |:|:|:|
 68 WLSAYS 74
 RESULT 36
 ADC95262
 ID ADC95262 standard; protein; 103 AA.
 XX
 AC ADC95262;
 XX

DT 01-JAN-2004 (first entry)
 XX
 DE E. faecium protein sequence SEQ ID 4889.
 XX
 KW Vaccine; urinary tract infection; bacteraemia; endocarditis; wound;
 KW abdominal-pelvic infection.
 XX
 OS Enterococcus faecium.
 PN US6583275-B1.
 PD 24-JUN-2003.
 XX
 PF 30-JUN-1998; 98US-00107532.
 XX
 PR 02-JUL-1997; 97US-0051571P.
 PR 14-MAY-1998; 98US-0085598P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 PI Doucette-Stamm LA, Bush D;
 XX
 DR WPI; 2003-799836/75.
 DR N-PSDB; ADC91608.
 XX
 PT New isolated nucleic acid derived from Enterococcus faecium encoding an
 PT Enterococcus faecium polypeptide useful for detection, prevention and
 PT treatment of a pathological condition resulting from a bacterial
 PT infection.
 XX
 PS Example 1; SEQ ID NO 4889; 243bp; English.
 XX
 CC The invention relates to an isolated nucleic acid derived from
 CC Enterococcus faecium encoding an Enterococcus faecium polypeptide having
 CC one of 10 fully defined sequences given in the (or comprising 40
 CC sequential nucleotides chosen from any of the nucleic acids, its
 CC complement or sequences hybridizing to it). Also included are a
 CC recombinant vector comprising the nucleic acid operably linked to
 CC transcription regulatory element, a cell comprising the vector and a
 CC single-stranded probe comprising the nucleic acid. The nucleic acids are
 CC chosen from 3654 disclosed sequences encoding 3654 disclosed proteins.
 CC The nucleic acids is useful for diagnosing pathological conditions
 CC resulting from E. faecium bacterial infection (e.g. urinary tract
 CC infection, bacteraemia, endocarditis, wounds and abdominal-pelvic
 CC infection) and for screening drugs such as agonists and antagonists. The
 CC nucleic acid is useful for recombinant production of candida albicans -
 CC derived peptides or antisense polypeptides. Pharmaceutical compositions
 CC and vaccines containing the nucleic acid are useful for preventing or
 CC treating Enterococcus faecium infections. The present sequence represents
 CC one if the disclosed E. faecium proteins.
 XX
 SQ Sequence 103 AA;
 QY
 Db Query Match 84.2%; Score 32; DB 7; Length 103;
 Best Local Similarity 83.3%; Pred. No. 1.4e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 1 WMLSAF 6
 |:|:|:|
 17 WMLSAF 22
 RESULT 37
 AAU25590
 ID AAU25590 standard; protein; 211 AA.
 XX
 AC AAU25590;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Human G Protein-Coupled Receptor (GPCR) polypeptide #37.
 XX
 KW Human; G-protein coupled receptor; GPCR; mental disorder; schizophrenia;

KW attention deficit disorder; anxiety; depression; bipolar disorder;
 KW neurological disorder; Huntington's disease; dementia; obesity; anorexia;
 KW metabolic disorder; Parkinson's disease; Tourette's syndrome; thrombosis;
 KW type 2 diabetes; cardiovascular disorder; myocardial infarction; cancer;
 KW cardiomyopathy; atherosclerosis; human immunodeficiency virus; HIV;
 KW viral infection; immunostimulant; neuroleptic; nootropic; tranquilizer;
 KW antidepressant; anorectic; gene therapy.
 XX Homo sapiens.
 XX
 PN WO200162797-A2.
 PD
 PD 30-AUG-2001.
 XX
 PF 23-FEB-2001; 2001WO-US005676.
 XX
 PR 23-FEB-2000; 2000US-0184247P.
 PR 23-FEB-2000; 2000US-0184303P.
 PR 23-FEB-2000; 2000US-0184304P.
 PR 23-FEB-2000; 2000US-0184305P.
 PR 23-FEB-2000; 2000US-0184397P.
 PR 02-MAR-2000; 2000US-0186457P.
 PR 03-MAR-2000; 2000US-0186810P.
 PR 03-MAR-2000; 2000US-0188064P.
 PR 13-MAR-2000; 2000US-0188860P.
 PR 03-APR-2000; 2000US-0194344P.
 PR 23-JUN-2000; 2000US-0213861P.
 PR 11-JUL-2000; 2000US-0217369P.
 PR 11-JUL-2000; 2000US-0217370P.
 PR 14-JUL-2000; 2000US-0218337P.
 PR 20-JUL-2000; 2000US-0218492P.
 XX
 PA (PNUA) PHARMACIA & UPJOHN CO.
 XX
 PI Vogel G, Wood LS, Parodi LA, Lind P;
 XX
 DR WPI; 2001-570628/64.
 DR N-PSDB; AAS42842.
 XX
 PT New isolated nucleic acid encoding a new G-protein coupled receptor
 PT polypeptide for detecting receptor modulators that can treat mental
 PT disorders, such as schizophrenia, anxiety, depression, or obesity.
 XX
 PS Claim 35; Page 84; 279pp; English.
 XX
 CC Sequences AAU25554-AAU25616 represent human G-protein coupled receptor
 CC (GPCR) polypeptides of the invention. The proteins and their associated
 CC DNA sequences can be used to identify compounds that modulate GPCR
 CC polypeptides and in screening for compounds that modulate GPCR activity.
 CC By screening a human subject for the presence of mutations in GPCR DNA, a
 CC GPCR-related disorder or a genetic predisposition can be diagnosed. The
 CC sequences can also be used for treatment and prevention of mental
 CC disorders such as schizophrenia, attention deficit disorder, anxiety,
 CC depression, dementia and bipolar disorder, neurological disorders such as
 CC Huntington's disease, Parkinson's disease and Tourette's syndrome,
 CC metabolic disorders such as obesity, anorexia and type 2 diabetes,
 CC cardiovascular disorders such as thrombosis, myocardial infarction,
 CC cardiomyopathy and atherosclerosis, viral infections caused by HIV and
 CC cancers
 XX
 SQ Sequence 211 AA;
 QY Query Match 84.2%; Score 32; DB 4; Length 211;
 Best Local Similarity 71.4%; Pred. No. 3.1e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 1 WMLSAFS 7
 128 WMLAAPS 134

RESULT 38
 ADR09699

ID ADR09699 standard; protein; 243 AA.
 XX
 XX ADR09699;
 XX
 DT 04-NOV-2004 (first entry)
 XX
 DE Human protein useful for treating neurological disease Seq 3205.
 XX
 XX human; oligo-capping method; diagnostic marker; gene therapy;
 KW osteoporosis; neurological disease; Alzheimer's disease;
 KW Parkinson's disease; dementia; short memory; cancer;
 KW sense or motor function; emotional reaction; fear response; panic;
 KW osteopathic; neuroprotective; nootropic; antiparkinsonian; cyostatic;
 KW tranquilizer.
 XX
 XX Homo sapiens.
 OS
 PN EPI447413-A2.
 XX
 PD 18-AUG-2004.
 XX
 PF 12-FEB-2004; 2004EP-00003145.
 XX
 PR 14-FEB-2003; 2003JP-00102207.
 PR 09-MAY-2003; 2003JP-00131452.
 XX
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.
 XX
 PI Isegai T, Yamamoto J, Nishikawa T, Isono Y, Sugiyama T, Otsuki T;
 PI Wakamatsu A, Ishii S, Nagai K, Irie R;
 XX
 DR WPI; 2004-583265/57.
 DR N-PSDB; ADR07743.
 XX
 PT New 1995 cDNA, useful for treating osteoporosis, neurological diseases,
 PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
 XX
 PS Claim 1; SEQ ID NO 3205; 2686pp; English.
 XX
 CC This invention relates to novel, isolated full length human cDNA
 CC molecules and the encoded proteins thereof. Specifically, it refers to
 CC cDNA clones obtained by an oligo-capping method, where none of these
 CC clones are identical to any known human mRNAs. The present invention
 CC describes an immunoassay to identify agonists and antagonists, as well as
 CC antibodies, antisense molecules and siRNAs that can all be used to bind
 CC to and modulate expression of the cDNA molecules. As such, these
 CC molecules are useful for diagnostic markers or therapeutic targets for
 CC the various diseases or morbid states. In particular, they are useful in
 CC gene therapy for treating osteoporosis, neurological disease, Alzheimer's
 CC disease, Parkinson's disease, dementia, short memory and various cancers,
 CC as well as for maintaining equilibrium of sense or motor function, and
 CC for treating emotional reaction, fear response and panic. Accordingly,
 CC they exhibit osteopathic, neuroprotective, nootropic, antiparkinsonian,
 CC cyostatic and tranquilizer activities. This polypeptide is a protein
 CC encoded by a full length human cDNA sequence of the invention. NOTE: This
 CC sequence is not given in the sequence listing of the specification but
 CC can be obtained on CD-ROM from the European Patent Office, Vienna Sub-
 CC office.
 XX
 SQ Sequence 243 AA;
 QY Query Match 84.2%; Score 32; DB 8; Length 243;
 Best Local Similarity 85.7%; Pred. No. 3.6e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 1 WMLSAFS 7
 187 WKLSAAPS 193

RESULT 39
 ABO73613
 ID ABO73613 standard; protein; 262 AA.

XX AB073613;
AC 29-JUL-2004. (first entry)
XX
XX Pseudomonas aeruginosa polypeptide #5788.
DE
XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
KM
XX Pseudomonas aeruginosa.
OS
XX US6551795-B1.
PN
XX 22-APR-2003.
PD
XX 18-FEB-1999; 99US-00252991.
PF
XX 18-FEB-1998; 98US-0074788P.
PR 27-JUL-1998; 98US-0094190P.
XX (GENO-) GENOME THERAPEUTICS CORP.
PA Rubenfield MJ, Nolling J, Deloughery C, Bush D;
PI WPI; 2003-615309/58.
XX N-PSDB; ABD07184.
DR
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
XX Disclosure; SEQ ID NO 22359; 455bp; English.
PS
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences AB067826-
CC AB084396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
XX Sequence 262 AA;
SQ
Query Match 84.2%; Score 32; DB 7; Length 262;
Best Local Similarity 85.7%; Pred. No. 3.9e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 1;
QY 1 WMLSAFS 7
Db 234 WMLSAIS 240

RESULT 40
AAG81234
ID AAG81234 standard; protein; 280 AA.
XX
XX AAG81234;
AC
XX 04-SEP-2001 (first entry)
DT
XX Mycobacterium tuberculosis potential drug target protein SEQ ID 283.
DE
XX Drug target; growth; organism viability; characterisation.
KM
XX

OS Mycobacterium tuberculosis.
XX
XX WO200135317-A1.
PN
XX 17-MAY-2001.
PD
XX 13-NOV-2000; 2000WO-US031152.
PF
XX 12-NOV-1999; 99US-0165086P.
PR 12-NOV-1999; 99US-0165124P.
XX 01-FEB-2000; 2000US-0179531P.
XX
XX (REGC) UNIV CALIFORNIA.
PA Eisenberg D, Rotstein SH, Marcotte EM;
XX WPI; 2001-329193/34.
PI
XX N-PSDB; AAH52085.
DR
XX Identifying nucleotide or polypeptide sequence for use as drug target,
PT involves providing algorithm that analyzes a functional relationship
PT between nucleotide or polypeptide sequences, and comparing the sequences.
XX
XX Disclosure; Page 190; 207pp; English.
PS
XX This invention relates to a method for identifying a nucleotide or
CC polypeptide sequence that may be a drug target, or essential for growth
CC or viability of an organism. Polynucleotide sequences AAH51947 - AAH52092
CC represent DNA encoding proteins AAG81096 - AAG81241, Mycobacterium
CC tuberculosis proteins which are potential drug targets. The DNA and
CC protein sequences are used to illustrate the method of the invention. The
CC method involves providing an unknown nucleotide or polypeptide sequences,
CC and comparing it to a number of sequences along with at least one
CC algorithm capable of analysing a functional relationship between
CC nucleotide and polypeptide sequences. The method is useful for
CC characterising the function of nucleic acids and polypeptides that may be
CC useful as a target for a drug or essential for the growth or viability of
CC an organism
XX
XX Sequence 280 AA;
SQ
Query Match 84.2%; Score 32; DB 4; Length 280;
Best Local Similarity 71.4%; Pred. No. 4.2e+02; Indels 0; Gaps 0;
Matches 5; Conservative 2; Mismatches 0;
QY 1 WMLSAFS 7
Db 262 WMLAAFA 268

RESULT 41
AAM51647
ID AAM51647 standard; protein; 289 AA.
XX
XX AAM51647;
AC
XX 20-FEB-2002 (first entry)
DT
XX Human GPCR polypeptide.
DE
XX Human, G-protein coupled receptor; GPCR; gene therapy;
KM MAS proto-oncogene receptor; human protease; disease.
XX Homo sapiens.
OS
XX WO200181409-A2.
PN
XX 01-NOV-2001.
PD
XX 24-APR-2001; 2001WO-US013097.
PF
XX 24-APR-2000; 2000US-0199149P.
PR 04-AUG-2000; 2000US-00633146.
XX

XX (PEKE) PE CORP NY.
 XX
 PI Wei M, Cravchik A, Di Francesco V, Beasley EM;
 XX
 DR WPI: 2002-0469265/06.
 DR N-PSDB; ABA03701.
 XX
 PT Novel human G protein-coupled receptor polypeptide that is related to MAS
 PT proto-oncogene receptor subfamily, useful as model and target for
 PT developing human therapeutic agent.
 XX
 PS Claim 1; Fig 2; 60pp; English.
 XX
 CC The invention relates to an isolated human G protein-coupled receptor
 CC (GPCR) polypeptide that is related to the MAS proto-oncogene receptor
 CC subfamily. The polypeptide comprises a fully defined sequence of 289
 CC amino acids as given in the specification, or its fragment comprising 10
 CC contiguous amino acids, or an amino acid sequence of an allelic variant
 CC or orthologue of the amino acid sequence given in the sequence. The
 CC polypeptide is useful for identifying a modulator of a GPCR polypeptide
 CC or an agent that binds to it. The polypeptide is also useful for
 CC treating a disease or condition mediated by human proteases. The present
 CC sequence is the polypeptide of the invention
 XX
 SQ Sequence 289 AA;
 Query Match 84.2%; Score 32; DB 5; Length 289;
 Best Local Similarity 71.4%; Pred. No. 4.3e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 WMLSAFS 7
 Db 94 WMLAAPS 100
 XX
 RESULT 42
 ADO29555 standard; protein; 289 AA.
 XX
 AC ADO29555;
 XX
 DT 29-JUL-2004 (first entry)
 XX
 DE Mouse GPCR N8 (MRGG) , SEQ ID NO:657.
 XX
 KW G protein-coupled receptor; GPCR; drug screening; diagnosis;
 KW transgenic mouse; neurological disorder; adrenal gland disorder;
 KW colon disorder; intestinal disorder; cardiovascular disorder;
 KW muscular disorder; blood disorder; immune disorder; bone disorder;
 KW joint disorder; metabolic disorder; nutritive disorder; cancer;
 KW kidney disorder; liver disorder; lung disorder; breast disorder;
 KW ovary disorder; uterus disorder; prostate disorder; testis disorder;
 KW skin disorder; stomach disorder; pancreas disorder; spleen disorder;
 KW thymus disorder; thyroid disorder; antiparkinsonian; antitumor;
 KW cytoskeletal; antiinflammatory; vasotropic; antiangiinal; antirhythmic;
 KW CNS; central nervous system; respiratory; antidiarrhoeic; antidiabetic;
 KW virucide; hepatotropic; antibacterial; antianemic; antiseborrhoeic;
 KW dermatological; antitumor; antithyroid; antiallergic; anorectic;
 KW immunosuppressive; nephrotropic; gene therapy; GPCR modulator; mouse;
 KW murine; receptor.
 XX
 OS Mus musculus.
 XX
 PN WO2004040000-A2.
 XX
 PD 13-MAY-2004.
 XX
 PF 09-SEP-2003; 2003WO-USO28226.
 XX
 PR 09-SEP-2002; 2002US-0409303P.
 PR 09-APR-2003; 2003US-0461329P.
 XX

PA (PRIM-) PRIMAL INC.
 XX
 PI Galtanaris GA, Bergmann JE, Gragerov A, Hohmann J, Li F,
 PI Madisen L, McIlwain KL, Pavlova MN, Vaessiliatis D, Zeng H;
 XX
 DR WPI: 2004-390329/36.
 DR N-PSDB; ADO30286.
 XX
 PT Novel mammalian G protein coupled receptors, useful for identifying
 PT compounds that modulates diagnosing and treating disease condition
 PT associated with GPCR dysfunction e.g. autoimmune diseases, angina
 PT pectoris, Parkinson's disease.
 XX
 PS Claim 151; SEQ ID NO 657; 542pp; English.
 XX
 CC The invention relates to human and mouse G protein-coupled receptors
 CC (GPCR) and nucleic acids encoding them. The invention also relates to
 CC sequences at least 90% identical to the GPCR proteins and nucleic acids
 CC of the invention; methods of treating, preventing or diagnosing diseases
 CC associated with GPCRs of the invention; methods of screening for
 CC compounds useful in the treatment of GPCR-related diseases; a transgenic
 CC mouse comprising a GPCR gene of the invention; a mouse comprising a
 CC mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived
 CC from the transgenic mice, kits comprising several mice, each of which has
 CC a mutation in a different GPCR gene of the invention; and kits comprising
 CC probes which hybridise to GPCR polynucleotides of the invention. The
 CC invention further discloses variants of the GPCR polypeptides and vectors
 CC comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may
 CC be used in the diagnosis, treatment or prevention of a wide variety of
 CC diseases including neurological disorders (e.g., Alzheimer's disease,
 CC depression, diabetic neuropathy, Parkinson's disease or schizophrenia);
 CC disorders of the adrenal gland; disorders of the colon or intestine
 CC (e.g., Crohn's disease, diarrhoea, food poisoning or irritable bowel
 CC syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or
 CC myocardial infarction); muscular disorders; blood disorders (e.g.,
 CC anaemia or leukaemia); immune disorders (e.g., autoimmune disorders or
 CC AIDS); bone and joint disorders (e.g., osteoarthritis, rheumatoid
 CC arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g.,
 CC obesity, enzyme deficiency-related diseases or vitamin deficiency-related
 CC diseases); and disorders of the kidney, liver, lung, breast, ovary,
 CC uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and
 CC thyroid (e.g., cancer). The present sequence represents a GPCR of the
 CC invention. Note: The full sequence data for this patent did not form part
 CC of the printed specification; those sequences not shown were obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 289 AA;
 Query Match 84.2%; Score 32; DB 8; Length 289;
 Best Local Similarity 71.4%; Pred. No. 4.3e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 WMLSAFS 7
 Db 94 WMLAAPS 100
 XX
 RESULT 43
 ADO29554 standard; protein; 289 AA.
 XX
 AC ADO29554;
 XX
 DT 29-JUL-2004 (first entry)
 XX
 DE Human GPCR N8 (MRGG) , SEQ ID NO:656.
 XX
 KW G protein-coupled receptor; GPCR; drug screening; diagnosis;
 KW transgenic mouse; neurological disorder; adrenal gland disorder;
 KW colon disorder; intestinal disorder; cardiovascular disorder;
 KW muscular disorder; blood disorder; immune disorder; bone disorder;
 KW joint disorder; metabolic disorder; nutritive disorder; cancer;
 KW

KW kidney disorder; liver disorder; lung disorder; breast disorder;
 KW ovary disorder; uterus disorder; prostate disorder; testis disorder;
 KW skin disorder; stomach disorder; pancreas disorder; spleen disorder;
 KW thymus disorder; thyroid disorder; antiparkinsonian; antianemic;
 KW cytoskeletal; antiinflammatory; vasorelaxant; antidiabetic;
 KW CNS; central nervous system; respiratory; antidiarrhoeal; antidiabetic;
 KW virucide; hepatotoxic; antibacterial; antianemic; antiseborrhoeic;
 KW dermatological; antitumor; anticholinergic; antiallergic; anorectic;
 KW immunosuppressive; nephrotoxic; gene therapy; GPCR modulator; human;
 KW receptor.
 KW Homo sapiens.
 OS
 XX
 XX WO2004040000-A2.
 PN
 XX 13-MAY-2004.
 PD
 XX 09-SEP-2003; 2003WO-US028226.
 PF
 XX 09-SEP-2002; 2002US-0409303P.
 PR
 XX 09-APR-2003; 2003US-0461329P.
 PA (PRIM-) PRIMAL INC.
 PI Galtnaris GA, Bergmann JE, Gragerov A, Hohmann J, Li F;
 PI Medisen L, McIlwain KL, Pavlova MN, Vassilakis D, Zeng H;
 XX
 XX WPI; 2004-390329/36.
 DR N-PDB; ADO29996.
 DR
 XX
 PT Novel mammalian G protein coupled receptors, useful for identifying
 PT compounds that modulates diagnosing and treating disease condition
 PT associated with GPCR dysfunction e.g. autoimmune diseases, angina
 PT pectoris, Parkinson's disease.
 PT
 XX
 XX Claim 151; SEQ ID NO 656; 542pp; English.
 PS
 XX
 XX The invention relates to human and mouse G protein-coupled receptors
 CC (GPCRs) and nucleic acids encoding them. The invention also relates to
 CC sequences at least 90% identical to the GPCR proteins and nucleic acids
 CC of the invention; methods of treating, preventing or diagnosing diseases
 CC associated with GPCRs of the invention; methods of screening for
 CC compounds useful in the treatment of GPCR-related diseases; a transgenic
 CC mouse comprising a GPCR gene of the invention; a mouse comprising a
 CC mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived
 CC from the transgenic mice; kits comprising several mice, each of which has
 CC a mutation in a different GPCR gene of the invention; and kits comprising
 CC probes which hybridize to GPCR polynucleotides of the invention. The
 CC invention further discloses variants of the GPCR polypeptides and vectors
 CC comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may
 CC be used in the diagnosis, treatment or prevention of a wide variety of
 CC diseases including neurological disorders (e.g., Alzheimer's disease,
 CC depression, diabetic neuropathy, Parkinson's disease or schizophrenia);
 CC disorders of the adrenal gland; disorders of the colon or intestine
 CC (e.g., Crohn's disease, diarrhoea, food poisoning or irritable bowel
 CC syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or
 CC myocardial infarction); muscular disorders; blood disorders (e.g.,
 CC anaemia or leukaemia); immune disorders (e.g., autoimmune disorders or
 CC AIDS); bone and joint disorders (e.g., osteoarthritis, rheumatoid
 CC arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g.,
 CC obesity, enzyme deficiency-related diseases or vitamin deficiency-related
 CC diseases); and disorders of the kidney, liver, lung, breast, ovary,
 CC uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and
 CC thyroid (e.g., cancers). The present sequence represents a GPCR of the
 CC invention. Note: The full sequence data for this patent did not form part
 CC of the printed specification; those sequences not shown were obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 XX Sequence 289 AA;
 SQ
 Query March 84.2%; Score 32; DB 8; Length 289;
 Best Local Similarity 71.4%; Pred. No. 4.3e+02;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 WMTSARS 7
 Db 94 WMTSARS 100
 RESULT 44
 AAU25612
 ID AAU25612 standard; protein; 323 AA.
 XX
 XX AAU25612;
 AC
 XX
 XX 18-DEC-2001 (first entry)
 DT
 XX
 XX Human G Protein-Coupled Receptor (GPCR) polypeptide #59.
 DE
 XX
 XX Human; G-protein coupled receptor; GPCR; mental disorder; schizophrenia;
 KW attention deficit disorder; anxiety; depression; bipolar disorder;
 KW neurological disorder; Huntington's disease; dementia; obesity; anorexia;
 KW metabolic disorder; Parkinson's disease; Tourette's syndrome; thrombosis;
 KW type 2 diabetes; cardiovascular disorder; myocardial infarction; cancer;
 KW cardiomyopathy; atherosclerosis; human immunodeficiency virus; HIV;
 KW viral infection; immunostimulant; neuroleptic; nootropic; tranquilliser;
 KW antidepressant; anorectic; gene therapy.
 KW
 XX
 XX Homo sapiens.
 OS
 XX
 XX WO200162797-A2.
 PN
 XX
 XX 30-AUG-2001.
 PD
 XX
 XX 23-FEB-2001; 2001WO-US005676.
 PF
 XX
 XX 23-FEB-2000; 2000US-0184247P.
 PR 23-FEB-2000; 2000US-0184303P.
 PR 23-FEB-2000; 2000US-0184304P.
 PR 23-FEB-2000; 2000US-0184305P.
 PR 23-FEB-2000; 2000US-0184397P.
 PR 02-MAR-2000; 2000US-0186457P.
 PR 03-MAR-2000; 2000US-0186810P.
 PR 09-MAR-2000; 2000US-0188064P.
 PR 13-MAR-2000; 2000US-0188880P.
 PR 03-APR-2000; 2000US-0194344P.
 PR 23-JUN-2000; 2000US-0213861P.
 PR 11-JUL-2000; 2000US-0217369P.
 PR 11-JUL-2000; 2000US-0217370P.
 PR 14-JUL-2000; 2000US-0218337P.
 PR 20-JUL-2000; 2000US-0218492P.
 XX
 XX (PHAA) PHARMACIA & UPJOHN CO.
 PA
 XX
 XX Vogel G, Wood LS, Parodi LA, Lind P;
 PI
 XX WPI; 2001-570628/64.
 DR N-PDB; AAS42864.
 DR
 XX
 XX New isolated nucleic acid encoding a new G-protein coupled receptor
 PT polypeptide for detecting receptor modulators that can treat mental
 PT disorders, such as schizophrenia, anxiety, depression, or obesity.
 PT
 XX
 XX Claim 35; Page 94; 279pp; English.
 PS
 XX
 XX Sequences AAU25554-AAU25616 represent human G-protein coupled receptor
 CC (GPCR) polypeptides of the invention. The proteins and their associated
 CC DNA sequences can be used to identify compounds which bind to GPCR
 CC polypeptides and in screening for compounds that modulate GPCR activity.
 CC By screening a human subject for the presence of mutations in GPCR DNA, a
 CC GPCR-related disorder or a genetic predisposition can be diagnosed. The
 CC sequences can also be used for treatment and prevention of mental
 CC disorders such as schizophrenia, attention deficit disorder, anxiety,
 CC depression, dementia and bipolar disorder, neurological disorders such as
 CC Huntington's disease, Parkinson's disease and Tourette's syndrome.

CC metabolic disorders such as obesity, anorexia and type 2 diabetes,
 CC cardiovascular disorders such as thrombosis, myocardial infarction,
 CC cardiomyopathy and atherosclerosis, viral infections caused by HIV and
 CC cancers

XX Sequence 323 AA;

Query Match 84.2%; Score 32; DB 4; Length 323;
 Best Local Similarity 71.4%; Pred. No. 4.9e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WMLSAFS 7
 Db 128 WMLAAPS 134

RESULT 45
 ABB06257
 ID ABB06257 standard; protein; 323 AA.

XX ABB06257;

DT 23-MAY-2002 (first entry)

DE Human G protein-coupled receptor TGR15 protein SEQ ID NO:1.

KM Human; G protein-coupled receptor; TGR15; nocotropic; antiinflammatory;
 KM vasotropic; immunomodulator; cytostatic; gene therapy; protein therapy;
 KM neurological; inflammatory; circulatory; degenerative; immune system;
 KM digestive disease; cancer.

XX Homo sapiens.

XX W0200204639-A1.

XX 17-JAN-2002.

XX 06-JUL-2001; 2001WO-JP005876.

XX 07-JUL-2000; 2000JP-00211987.

XX 26-DEC-2000; 2000JP-00395566.

XX (TAKE) TAKEDA CHEM IND LTD.

XX Moriya T, Ito T, Shintani Y, Miyajima N;

XX WPI: 2002-179705/23.

XX N-PSDB; ABL40211.

PT G-protein coupled receptor protein TGR15 of human origin and DNA encoding
 PT it for diagnosis and treatment of cancer and circulatory and other
 PT diseases associated with its expression.

XX Claim 1; Fig 2; 103pp; Japanese.

CC The present sequence represents a human guanine nucleotide binding
 CC protein (G protein)-coupled receptor protein designated TGR15. TGR15 has
 CC nocotropic, antiinflammatory, vasotropic, immunomodulator and cytostatic
 CC activities. The TGR15 polynucleotide and protein can be used in gene
 CC therapy and protein therapy. G-protein coupled receptor proteins are cell
 CC membrane proteins which mediate the cellular response to a large variety
 CC of signalling molecules. The TGR15 polynucleotide and protein can be used
 CC in the diagnosis, treatment and prevention of diseases including
 CC neurological, inflammatory, circulatory, degenerative, immune system and
 CC digestive diseases and cancer

XX Sequence 323 AA;

Query Match 84.2%; Score 32; DB 5; Length 323;
 Best Local Similarity 71.4%; Pred. No. 4.9e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WMLSAFS 7

Db 128 WMLAAPS 134

RESULT 46
 ABB10905
 ID ABB10905 standard; protein; 371 AA.

XX ABB10905;

DT 05-DEC-2002 (first entry)

DE Human secreted protein (SECP) #1.

KM Human; gene therapy; secreted protein; SECP; hepatitis; cancer;
 KM cell proliferative disorder; autoimmune disorder; inflammatory disorder;
 KM AIDS; asthma; anaemia; allergy; atopic dermatitis; myocardial infarction;
 KM cardiovascular disorder; vascular tumour; neurological disorder; stroke;
 KM epilepsy; cerebral neoplasm; Alzheimer's disease; developmental disorder;
 KM Cushing's syndrome; muscular dystrophy.

XX Homo sapiens.

XX W0200270669-A2.

XX 12-SEP-2002.

XX 05-MAR-2002; 2002WO-US007719.

XX 06-MAR-2001; 2001US-0273946P.

XX 16-MAR-2001; 2001US-0276873P.

XX 30-MAR-2001; 2001US-0280531P.

XX 30-MAR-2001; 2001US-0280596P.

XX 16-NOV-2001; 2001US-0332426P.

XX 28-NOV-2001; 2001US-0334229P.

XX 11-JAN-2002; 2002US-0347703P.

XX (INCY-) INCYTE GENOMICS INC.

XX WPI: 2002-713444/77.

XX N-PSDB; ABL11179.

PT New human secreted proteins and nucleic acids useful in diagnosing,
 PT treating and preventing cell proliferative, autoimmune/inflammatory,
 PT cardiovascular, neurological, and developmental disorders.

XX Claim 1; Page 126-127; 162pp; English.

CC The invention comprises the amino acid and coding sequences of human
 CC secreted proteins (SECP). The human SECP DNA and protein sequences of the
 CC invention are useful for the treatment and prevention of cell
 CC proliferative disorders (e.g. actinic keratosis, arteriosclerosis,
 CC bursitis, hepatitis or cancer); autoimmune/inflammatory disorders (e.g.
 CC AIDS, asthma, anaemia, allergies or atopic dermatitis); cardiovascular
 CC disorders (e.g. congestive heart failure, ischaemic heart disease,
 CC myocardial infarction, hypertensive heart disease, or vascular tumours);
 CC neurological disorders (e.g. epilepsy, stroke, cerebral neoplasms, or
 CC Alzheimer's disease); and developmental disorders (e.g. renal tubular
 CC acidosis, Cushing's syndrome, Duchenne and Becker muscular dystrophy, or
 CC hypothyroidism). The present amino acid sequence represents a human
 CC secreted protein (SECP) of the invention

XX Sequence 371 AA;

Query Match 84.2%; Score 32; DB 5; Length 371;
 Best Local Similarity 85.7%; Pred. No. 5.7e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY      1 WMLSAFS 7
      | | | | |
Db      187 WMLSAFS 193

RESULT 47
ABO60886
ID      ABO60886 standard; protein; 437 AA.
XX
XX      ABO60886;
AC
XX
XX      29-JUL-2004 (first entry)
DT
XX
XX      Klebsiella pneumoniae polypeptide seqid 7403.
DE
XX      Recombinant expression vector; transcription regulatory element;
KW      Klebsiella pneumoniae protein; antibacterial; Vaccine.
XX
XX      Klebsiella pneumoniae.
OS
XX      US610836-B1.
PN
XX      26-AUG-2003.
PD
XX      27-JAN-2000; 2000US-00489039.
PF
XX      29-JAN-1999; 99US-0117747P.
PR
XX      (GENO-) GENOME THERAPEUTICS CORP.
PA
XX      Breton GU, Osborne M;
PI
XX      WPI; 2003-895346/82.
XX      DR
XX      N-PSDB; ACH94437.
DR
XX
XX      New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
PT      preparing a vaccine composition against Klebsiella pneumoniae.
XX
XX      Disclosure; SEQ ID NO 7403; 932pp; English.
PS
XX
XX      The invention describes a new isolated nucleic acid encoding a Klebsiella
CC      pneumoniae polypeptide. Also described are: a recombinant expression
CC      vector comprising the nucleic acid, operably linked to a transcription
CC      regulatory element; and a cell comprising the recombinant expression
CC      vector. The nucleic acid is useful for preparing a vaccine composition
CC      against Klebsiella pneumoniae. This is the amino acid sequence of a
CC      Klebsiella pneumoniae polypeptide of the invention
XX
XX      Sequence 437 AA;
SQ

Query Match      84.2%; Score 32; DB 7; Length 437;
Best Local Similarity 85.7%; Pred. No. 6.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY      1 WMLSAFS 7
      | | | | |
Db      391 WMLSAFS 397

RESULT 48
ADF70484
ID      ADF70484 standard; protein; 561 AA.
XX
XX      ADF70484;
AC
XX
XX      12-FEB-2004 (first entry)
DT
XX
XX      Orphan receptor ligand-related human protein SeqID107.
DE
XX      ligand; orphan receptor protein; fusion protein; fluorescent protein;
KW      cell expression; green fluorescent protein; GFP; GFP-1; wild-type GFP;
KW      GFPuv; Enhanced GFP; EGFP; human.
XX
XX

```

```

OS      Homo sapiens.
XX
XX      MO2003071272-A1.
PN
XX      28-AUG-2003.
PD
XX      21-FEB-2003; 2003WO-JP001901.
XX      PF
XX      22-FEB-2002; 2002JP-00045728.
PR      23-JUL-2002; 2002JP-0021949.
XX      PR      11-OCT-2002; 2002JP-00298237.
XX      (TAKE ) TAKEDA CHEM IND LTD.
PA
XX      Hinuma S, Fujii R, Ogi K, Komatsu H, Kawamata Y, Hosoya M;
PI      WPI; 2003-697654/66.
XX      DR      N-PSDB; ADF70586.
DR
XX      Transformation of cells with a fusion protein of an orphan receptor
PT      protein with a fluorescent protein useful for identification of ligands
XX      to the orphan receptor.
XX
XX      Disclosure; SEQ ID NO 107; 594pp; Japanese.
PS
XX
XX      This invention relates to a novel method of identifying ligands to an
CC      orphan receptor protein which comprises transforming cells with DNA
CC      encoding a fusion protein of the orphan receptor with a fluorescent
CC      protein, so that the fusion protein is expressed in the cells (or cell
CC      membranes isolated from them) and contacting the cells with the potential
CC      ligand to be tested. A suitable fluorescent protein for incorporation in
CC      the fusion protein is green fluorescent protein (GFP), for example GFP-1,
CC      wild-type GFP, GFPuv or Enhanced GFP (EGFP). The method is useful for the
CC      identification of ligands binding to an orphan receptor protein.
XX
XX      Sequence 561 AA;
SQ

Query Match      84.2%; Score 32; DB 7; Length 561;
Best Local Similarity 71.4%; Pred. No. 8.8e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY      1 WMLSAFS 7
      | | | | |
Db      128 WMLSAFS 134

RESULT 49
ADU02401
ID      ADU02401 standard; protein; 1180 AA.
XX
XX      ADU02401;
AC
XX      27-JAN-2005 (first entry)
DT
XX
XX      Novel human polypeptide seqid 868.
DE
XX      cytosstatic; antipsoiatric; antiinflammatory; gene therapy; Nanodisc;
KW      proliferative disorder; inflammatory disorder; immune disorder;
KW      metabolic disorder; bone disorder; CNS disorder; cancer; psoriasis;
KW      ulcerative colitis; human.
XX
XX      Homo sapiens.
OS
XX      WO2004093804-A2.
XX      PN
XX      04-NOV-2004.
XX      PD
XX      19-APR-2004; 2004WO-US012047.
XX      PF
XX      18-APR-2003; 2003US-0463708P.
XX      PR      18-APR-2003; 2003US-0463732P.
XX      PR      02-MAY-2003; 2003US-0467199P.
XX      PR      02-MAY-2003; 2003US-0467230P.

```

PR 19-MAY-2003; 2003US-0471306P.
 PR 19-MAY-2003; 2003US-0471336P.
 PR 08-JUL-2003; 2003US-0485223P.
 PR 08-JUL-2003; 2003US-0485224P.
 PR 14-JUL-2003; 2003US-0486446P.
 PR 14-JUL-2003; 2003US-0486480P.
 PR 08-AUG-2003; 2003US-0493573P.
 PR 08-AUG-2003; 2003US-0493577P.
 PR 08-SEP-2003; 2003US-0505059P.
 XX
 PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
 XX
 PI Lee E, Hestir K, Chu K, Masuoka L, Williams LT;
 XX
 DR WPI; 2004-775861/76.
 DR N-PSDB; ADU01669.
 XX
 PT New first nucleic acid molecule comprising a polynucleotide sequence
 PT diagnosing or treating e.g., cancer, psoriasis or ulcerative colitis.
 XX
 PS Claim 14; SEQ ID NO 868; 291bp; English.
 XX
 CC The invention describes a new first nucleic acid molecule comprising a
 CC polynucleotide sequence given in the specification. Also described are:
 CC an animal injected with the nucleic acid molecule; a second nucleic acid
 CC molecule comprising a second polynucleotide sequence that is at least
 CC about 70, 80, 90 or 95% homologous to the first nucleic acid molecule or
 CC that hybridizes to the first polynucleotide sequence under high
 CC stringency conditions; a vector comprising the nucleic acid molecule and
 CC a promoter that drives the expression of the nucleic acid molecule; a
 CC host cell transformed, transfected, transduced or infected with the
 CC nucleic acid molecule; a nucleic acid composition comprising a carrier or
 CC a buffer and one or more compositions comprising the nucleic acid
 CC molecule, vector or host cell; a substantially purified polypeptide; an
 CC animal injected with the polypeptide; a polypeptide composition
 CC comprising the polypeptide molecule and a carrier or buffer; a cell
 CC culture medium comprising the polypeptide or transfected cells
 CC transfected with the polynucleotide; making a transformed, transfected,
 CC transduced, or infected host cell; synthesizing Nanodiscs simultaneously
 CC and for synthesising a series of simultaneously-synthesised Nanodiscs
 CC sequentially utilising a dynamic system; preparing a hydrophobic protein
 CC for determination of crystal structure; immunising a non-human animal;
 CC screening for modulators of hydrophobic protein activity; a diagnostic
 CC kit; determining the presence of the nucleic acid molecule or its
 CC complement; determining the presence of an antibody to the polypeptide in
 CC a sample; an antibody specifically recognising, binding to or modulating
 CC the biological activity of at least one polypeptide encoded by a nucleic
 CC acid molecule or its biologically active fragment; an antibody
 CC composition comprising the antibody and a carrier; a bacteriophage, where
 CC the antibody is displayed on the bacteriophage; a bacterial cell
 CC comprising the bacteriophage; a non-human animal injected with the
 CC antibody composition; a host cell that secretes the antibody; making an
 CC antibody; diagnosing a disease, disorder, syndrome, or condition
 CC comprising cancer, or proliferative, inflammatory, immune, metabolic,
 CC bone, CNS, genetic, bacterial and viral diseases, disorders, syndromes or
 CC conditions in a patient; a modulator composition comprising a modulator
 CC and a carrier; gene therapy; prophylactic or therapeutic treatment of a
 CC subject; an isolated modified cell comprising at least one first
 CC heterologous nucleic acid molecule, where the first heterologous nucleic
 CC acid molecule comprises a first polynucleotide sequence that encodes a
 CC first polypeptide; a non-human animal deficient in the polypeptide or
 CC that over-expresses the polypeptide; isolated tissues derived from the
 CC non-human animal; and one or more cells derived from the non-human
 CC animal. The nucleic acid is useful in preparing a composition for
 CC diagnosing or treating e.g., cancer, psoriasis or ulcerative colitis.
 CC This is the amino acid sequence of a novel human polypeptide of the
 CC invention.
 XX
 SQ Sequence 1180 AA;
 Query Match 84.2%; Score 32; DB 8; Length 1180;
 Best Local Similarity 71.4%; Pred. No. 2e+03;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 WMLSAFS 7
 Db 746 WMLSAFS 752
 RESULT 50
 ABG21342
 ID ABG21342 standard; protein; 1923 AA.
 XX
 AC ABG21342;
 XX
 DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #21333.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN W0200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US008631.
 XX
 PR 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX
 PA (HYSE-) HYSEQ INC.
 PI Dmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR N-PSDB; AAS85529.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 20; SEQ ID NO 51701; 103bp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantifying a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC entry did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPD at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 1923 AA;
 Query Match 84.2%; Score 32; DB 4; Length 1923;
 Best Local Similarity 71.4%; Pred. No. 3.3e+03;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 1 WMLSAFS 7
|:|:|
Db 208 WMLAAPS 214

Search completed: August 29, 2006, 05:59:27
Job time : 68.7426 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2006, 06:00:58 ; Search time 7 Seconds
(without alignments)
96.217 Million cell updates/sec

Title: US-10-541-343-3

Perfect score: 38

Sequence: 1 WMLSAFS 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database :

PIR 80:*
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	34	89.5	637	2	A56263	beta-galactosidase
2	32	84.2	187	2	T30536	hypothetical prote
3	32	84.2	280	2	C70696	probable transpor
4	32	84.2	345	2	AB1813	hypothetical prote
5	32	84.2	392	2	D83934	hypothetical prote
6	32	84.2	465	1	W2BBB4	51.3K capsid prote
7	31	81.6	149	2	D84181	hypothetical prote
8	31	81.6	187	2	T51719	hypothetical prote
9	31	81.6	257	2	H83697	hypothetical prote
10	31	81.6	265	2	T01568	hypothetical prote
11	31	81.6	365	2	E97674	hypothetical prote
12	31	81.6	368	2	AD2899	conserved hypochet
13	31	81.6	371	1	OPUG65	diarylpropane pero
14	31	81.6	371	1	JC1268	hypothetical prote
15	31	81.6	372	1	OPUG3P	hypothetical prote
16	31	81.6	372	1	OPUGAP	hypothetical prote
17	31	81.6	372	1	OPUGBP	hypothetical prote
18	31	81.6	372	2	S01028	hypothetical prote
19	31	81.6	372	2	A36693	hypothetical prote
20	31	81.6	372	2	JT0402	hypothetical prote
21	31	81.6	372	2	A43638	hypothetical prote
22	31	81.6	372	2	S69246	hypothetical prote
23	31	81.6	372	2	B43638	hypothetical prote
24	31	81.6	372	2	A32322	hypothetical prote
25	31	81.6	372	2	US0065	hypothetical prote
26	31	81.6	372	2	B32322	hypothetical prote
27	31	81.6	372	2	JH0156	hypothetical prote
28	31	81.6	373	2	A27817	hypothetical prote
29	31	81.6	416	2	D96955	sensory transducti

30	31	81.6	541	2	T48836	hypothetical prote
31	31	81.6	1121	2	C82120	transcription regu
32	30	78.9	182	2	B53189	cytochrome b560 ho
33	30	78.9	271	2	T11590	hypothetical prote
34	30	78.9	305	2	D96769	hypothetical prote
35	30	78.9	332	2	G64141	probable tetracycl
36	30	78.9	341	2	C90322	glycosyltransferas
37	30	78.9	348	2	B97444	tetracyclidsacchar
38	30	78.9	348	2	AC2662	tetracyclidsacchar
39	30	78.9	374	2	E82168	probable alpha-1,6
40	30	78.9	400	2	D82419	tyrosine-specific
41	30	78.9	401	2	S65138	glycoprotein antiq
42	30	78.9	409	2	T11743	p47 protein - pig
43	30	78.9	412	2	F95394	probable transmemb
44	30	78.9	417	2	AC0546	probable metabolit
45	30	78.9	427	2	S74211	PAS-6/7 protein pr
46	30	78.9	454	2	A75444	hypothetical prote
47	30	78.9	467	2	F90544	conserved hypochet
48	30	78.9	501	2	S70171	integral membrane
49	30	78.9	548	2	T40598	hypothetical prote
50	30	78.9	906	2	A39280	glutamate receptor
51	30	78.9	925	2	T01384	hypothetical prote
52	30	78.9	1321	2	T29308	hypothetical prote
53	30	78.9	1372	2	T29309	hypothetical prote
54	29	76.3	119	2	E95955	hypothetical prote
55	29	76.3	158	2	B82190	hypothetical prote
56	29	76.3	201	2	AD2503	hypothetical prote
57	29	76.3	220	2	S40931	hypothetical prote
58	29	76.3	252	2	F69955	hypothetical prote
59	29	76.3	279	2	H86922	probable ABC trans
60	29	76.3	279	2	C64620	hypothetical prote
61	29	76.3	279	2	B71894	hypothetical prote
62	29	76.3	301	1	S12864	retinal isomerase
63	29	76.3	302	2	S60777	citg protein - Kle
64	29	76.3	313	2	S47433	cathapsin L (EC 3.
65	29	76.3	320	2	S19651	cysteine proteinas
66	29	76.3	329	2	H83252	glycosyltransferas
67	29	76.3	347	2	D84649	hypothetical prote
68	29	76.3	350	2	T05349	hypothetical prote
69	29	76.3	352	2	B34284	NADH dehydrogenas
70	29	76.3	356	2	A11393	conserved hypochet
71	29	76.3	390	2	D69757	multidrug-efflux t
72	29	76.3	411	2	AB0013	probable exported
73	29	76.3	413	1	IKEC5A	colicin V secretio
74	29	76.3	431	2	H70731	probable esterase
75	29	76.3	432	2	I50829	alpha 2-adrenocept
76	29	76.3	446	2	T45358	two-component sens
77	29	76.3	446	2	C70783	hypothetical prote
78	29	76.3	450	2	T21515	hypothetical prote
79	29	76.3	479	2	T50726	hypothetical prote
80	29	76.3	514	2	S54019	hypothetical prote
81	29	76.3	529	2	G75577	extracellular solu
82	29	76.3	553	2	AC2931	hypothetical prote
83	29	76.3	564	2	G86358	protein similar to
84	29	76.3	565	2	B86359	ABC transporter AT
85	29	76.3	608	2	AD2000	hypothetical prote
86	29	76.3	611	2	T04510	hypothetical prote
87	29	76.3	630	2	C98351	oligopeptide-bindi
88	29	76.3	641	2	AG2164	hypothetical prote
89	29	76.3	685	2	S65974	conserved hypochet
90	29	76.3	759	2	B90520	ABC transporter at
91	29	76.3	767	2	T22925	hypothetical prote
92	29	76.3	767	2	E96813	probable Oxidodqua
93	29	76.3	933	1	BVBCCC	sensor protein rcs
94	29	76.3	933	2	C91017	sensor for ctr cap
95	29	76.3	933	2	E85861	hypothetical prote
96	29	76.3	954	2	S60301	H+-exporting ATPas
97	29	76.3	954	2	H96838	hypothetical prote
98	29	76.3	1065	2	A70797	hypothetical prote
99	29	76.3	1151	2	T33777	hypothetical prote
100	29	76.3	4063	2	T42893	probable spectrin
101	29	76.3	4101	2	T36360	hypothetical prote
102	28	73.7	65	2	T36041	hypothetical prote

103	28	73.7	71	2	A64629	176	28	73.7	533	2	S18539	actva-1 protein -
104	28	73.7	71	2	D71885	177	28	73.7	550	2	A82378	conserved hypotet
105	28	73.7	79	2	AD3009	178	28	73.7	560	2	S46734	hypothetical prote
106	28	73.7	85	2	B98275	179	28	73.7	584	1	VCMS7A	env polyprotein pr
107	28	73.7	97	2	CG9841	180	28	73.7	616	2	E84424	probable auxin tra
108	28	73.7	104	2	F81814	181	28	73.7	647	2	T51808	probable auxin eft
109	28	73.7	137	2	D75337	182	28	73.7	663	2	T30621	hypothetical prote
110	28	73.7	153	2	G69660	183	28	73.7	716	2	D69855	conserved hypotet
111	28	73.7	162	2	C32322	184	28	73.7	722	2	F87662	conserved hypotet
112	28	73.7	162	2	A10884	185	28	73.7	830	2	S50810	FeonB-dependent rec
113	28	73.7	162	2	B91113	186	28	73.7	843	2	A40970	undulin 1 - human
114	28	73.7	162	2	B85958	187	28	73.7	880	2	T02245	hypothetical prote
115	28	73.7	162	2	F65085	188	28	73.7	884	2	S55651	DNA helicase-prima
116	28	73.7	163	2	D64025	189	28	73.7	889	2	JC6015	chitin synthase (E
117	28	73.7	174	2	D91117	190	28	73.7	1061	2	D98008	conserved hypotet
118	28	73.7	188	2	S74963	191	28	73.7	1085	2	H82511	hypothetical prote
119	28	73.7	194	2	A11038	192	28	73.7	1085	2	S40476	Ca(2+)-sensing rec
120	28	73.7	207	2	S53801	193	28	73.7	1116	2	S41915	DNA-directed RNA p
121	28	73.7	209	2	E82388	194	28	73.7	1136	2	T40355	hypothetical prote
122	28	73.7	212	2	PS0012	195	28	73.7	1136	2	VJCH2	vitellogenin II pr
123	28	73.7	213	2	C82953	196	28	73.7	1852	1	VJCH2	probable non-ribos
124	28	73.7	263	2	A21195	197	27	71.1	106	2	H83343	hypothetical prote
125	28	73.7	266	2	UC7300	198	27	71.1	110	2	AC1958	hypothetical prote
126	28	73.7	267	2	UC4857	199	27	71.1	145	2	H69618	stresses- and starva
127	28	73.7	288	2	S17659	200	27	71.1	148	2	E82925	hypothetical prote
128	28	73.7	291	2	B85018	201	27	71.1	148	2	E84023	hypothetical prote
129	28	73.7	299	2	B82504	202	27	71.1	162	2	F84119	ATP synthase subun
130	28	73.7	302	2	C96531	203	27	71.1	166	2	AG3566	conserved hypotet
131	28	73.7	305	2	T52111	204	27	71.1	174	2	F83514	protein-disulfide
132	28	73.7	305	2	B84848	205	27	71.1	176	1	F64864	disulfide bond for
133	28	73.7	307	2	T52109	206	27	71.1	176	2	AH0723	hypothetical prote
134	28	73.7	308	2	T46026	207	27	71.1	176	2	H85696	protein-disulfide
135	28	73.7	311	2	G84719	208	27	71.1	177	2	H90838	probable protein-d
136	28	73.7	313	2	T11272	209	27	71.1	195	2	S24997	formate C-acetyltr
137	28	73.7	315	2	AB1812	210	27	71.1	199	2	T48763	hypothetical prote
138	28	73.7	315	2	T50562	211	27	71.1	219	2	A70314	deoxyribose-phosph
139	28	73.7	315	2	T50561	212	27	71.1	227	2	B37206	class II histocomp
140	28	73.7	318	2	T20063	213	27	71.1	227	2	A47635	MHC class II histo
141	28	73.7	321	2	T29419	214	27	71.1	227	2	G97310	tumor necrosis fac
142	28	73.7	326	2	AE1169	215	27	71.1	234	1	JH0529	hypothetical prote
143	28	73.7	335	2	A10295	216	27	71.1	265	2	E84533	probable membrane
144	28	73.7	345	2	T12347	217	27	71.1	265	2	T46952	hypothetical prote
145	28	73.7	347	2	H69867	218	27	71.1	266	2	AF1617	hypothetical prote
146	28	73.7	357	2	AD1769	219	27	71.1	266	2	AH1254	fibronectin-bindin
147	28	73.7	357	2	SS9531	220	27	71.1	274	1	A41461	hypothetical prote
148	28	73.7	372	1	OBJGH2	221	27	71.1	277	2	T31855	conserved hypotet
149	28	73.7	380	2	T01706	222	27	71.1	280	2	H69288	capsular polysacch
150	28	73.7	383	2	H00482	223	27	71.1	282	2	T50040	hypothetical prote
151	28	73.7	385	2	S68780	224	27	71.1	290	2	E72676	dehydrogenase rela
152	28	73.7	387	2	B83553	225	27	71.1	292	1	CRH05	carbonate dehydrat
153	28	73.7	395	2	B96610	226	27	71.1	305	1	G97050	DNA-methyltransfer
154	28	73.7	400	2	C70082	227	27	71.1	314	2	T47971	seven in absentia-
155	28	73.7	404	2	H83249	228	27	71.1	315	2	T27372	hypothetical prote
156	28	73.7	408	2	B84591	229	27	71.1	320	2	T27372	movement protein -
157	28	73.7	413	2	B83180	230	27	71.1	322	2	S45000	ribonucleoside-dip
158	28	73.7	426	2	T32206	231	27	71.1	324	2	B71655	hypothetical prote
159	28	73.7	429	2	E84410	232	27	71.1	327	2	T09027	hypothetical prote
160	28	73.7	429	2	T08562	233	27	71.1	328	2	C97781	DNA polymerase III
161	28	73.7	439	2	B75487	234	27	71.1	330	2	C82777	probable integral
162	28	73.7	440	2	F83235	235	27	71.1	334	2	B81406	tryptophan-tRNA II
163	28	73.7	440	2	S11793	236	27	71.1	335	2	E84992	SINAH1 protein (Im
164	28	73.7	442	2	T16773	237	27	71.1	336	2	T50560	hypothetical prote
165	28	73.7	449	2	D85962	238	27	71.1	340	2	B86363	hypothetical prote
166	28	73.7	449	2	H65089	239	27	71.1	350	2	T25156	hypothetical prote
167	28	73.7	466	2	H81697	240	27	71.1	364	2	T25155	hypothetical prote
168	28	73.7	466	2	G71542	241	27	71.1	374	2	T33173	hypothetical prote
169	28	73.7	469	2	D70048	242	27	71.1	388	2	G90450	probable membrane
170	28	73.7	483	2	F81221	243	27	71.1	393	2	S49759	hypothetical prote
171	28	73.7	490	2	S77201	244	27	71.1	397	2	C84078	porin O precursor
172	28	73.7	494	2	H81992	245	27	71.1	398	2	A82820	hypothetical prote
173	28	73.7	511	2	S44275	246	27	71.1	399	2	T21586	hypothetical prote
174	28	73.7	513	2	S63701	247	27	71.1	399	2	T27853	hypothetical prote
175	28	73.7	521	2	D96992	248	27	71.1	417	2	T34561	hypothetical prote

249	27	71.1	423	2	D89949	folypolyglutamate
250	27	71.1	435	2	S78258	probable transloca
251	27	71.1	426	2	T22568	hypothetical prote
252	27	71.1	431	2	A95419	probable transmem
253	27	71.1	434	2	AF0683	hexonate transpo
254	27	71.1	438	2	T39268	hypothetical prote
255	27	71.1	440	1	R6ECUC	regulatory protein
256	27	71.1	440	2	F86050	regulator of ubpt
257	27	71.1	440	2	D91204	regulator of ubpt
258	27	71.1	443	2	AG0487	probable regulator
259	27	71.1	444	2	AD0238	heaman storage sys
260	27	71.1	450	2	T40446	metaxin homolog -
261	27	71.1	457	2	T47005	hypothetical prote
262	27	71.1	459	2	F70884	probable transposa
263	27	71.1	471	2	AE0838	probable two-compo
264	27	71.1	483	2	A11908	hypothetical prote
265	27	71.1	493	2	T48630	high affinity nitr
266	27	71.1	500	2	S49302	AMU1218 protein -
267	27	71.1	502	2	C86263	hypothetical prote
268	27	71.1	503	2	B84007	cardiolipin synthe
269	27	71.1	509	1	O4RTLO	laureate omega-hydr
270	27	71.1	509	2	S47553	cytochrome P450 Cy
271	27	71.1	529	2	T23190	hypothetical prote
272	27	71.1	537	2	T38015	hypothetical prote
273	27	71.1	575	2	S50962	hypothetical prote
274	27	71.1	577	2	A53202	cyclophilin C-asso
275	27	71.1	583	2	G84629	probable PTR2 fam1
276	27	71.1	595	2	B64460	hypothetical prote
277	27	71.1	610	2	A84417	hypothetical prote
278	27	71.1	610	2	D83656	hypothetical prote
279	27	71.1	633	2	S76749	hypothetical prote
280	27	71.1	635	2	T41367	hypothetical prote
281	27	71.1	643	2	S17997	hypothetical gene
282	27	71.1	677	2	T08943	hypothetical prote
283	27	71.1	679	2	G83492	hypothetical prote
284	27	71.1	701	2	C84972	hypothetical prote
285	27	71.1	706	2	AD0173	conserved hypochet
286	27	71.1	794	2	T23048	hypothetical prote
287	27	71.1	918	2	S45872	hypothetical prote
288	27	71.1	940	2	T48700	hypothetical prote
289	27	71.1	1097	2	S17308	leukemia inhibitor
290	27	71.1	1164	2	S46769	hypothetical prote
291	27	71.1	1179	2	T05673	hypothetical prote
292	27	71.1	1245	2	G84897	hypothetical prote
293	27	71.1	1257	2	T28937	hypothetical prote
294	27	71.1	1341	2	T18301	lactrophilin-2, sp1
295	27	71.1	1354	2	T18375	lactrophilin-2, sp1
296	27	71.1	1356	2	T18367	lactrophilin-2, sp1
297	27	71.1	1369	2	T18379	lactrophilin-2, sp1
298	27	71.1	1379	2	T13718	poliux gene protei
299	27	71.1	1384	2	T18366	lactrophilin-2, sp1
300	27	71.1	1397	2	T18377	lactrophilin-2, sp1
301	27	71.1	1397	2	T18370	lactrophilin-2, sp1
302	27	71.1	1407	2	T18381	lactrophilin-2, sp1
303	27	71.1	1412	2	T18380	lactrophilin-2, sp1
304	27	71.1	1420	2	T18385	lactrophilin-2, sp1
305	27	71.1	1420	2	T17158	CU2AB protein - ra
306	27	71.1	1422	2	T18383	bullous pemphigoid
307	27	71.1	1433	2	A46053	bullous pemphigoid
308	27	71.1	1435	2	T18387	lactrophilin-2, sp1
309	27	71.1	1435	2	T46611	CU2BB protein - ra
310	27	71.1	1439	2	T02087	gag/pol polyprotei
311	27	71.1	1450	2	T18382	lactrophilin-2, sp1
312	27	71.1	1452	2	T17157	CU2AA protein - ra
313	27	71.1	1463	2	T18386	lactrophilin-2, sp1
314	27	71.1	1463	2	T17159	CU2AC protein - ra
315	27	71.1	1465	2	T18384	lactrophilin-2, sp1
316	27	71.1	1467	2	T17150	CU2BA protein - ra
317	27	71.1	1468	2	T05672	hypothetical prote
318	27	71.1	1478	2	T18388	lactrophilin-2, sp1
319	27	71.1	1478	2	T17185	CU2BC protein - ra
320	27	71.1	1487	2	T14324	alpha-lactotoxin r
321	27	71.1	1680	2	T41628	probable transcrip
322	27	71.1	1703	2	S15047	SNF2 protein - yea
323	27	71.1	1784	2	T02844	cdc16-related prot
324	27	71.1	1905	2	T18267	multidrug resistan
325	27	71.1	3844	2	T18402	asparagine/asparta
326	27	71.1	4085	2	S28600	apolipoprotein C-I
327	27	71.1	66	2	S13188	apolipoprotein C-I
328	26	68.4	69	2	H97817	hypothetical prote
329	26	68.4	72	1	MSW16B	E5B protein - huma
330	26	68.4	72	2	D61055	E5b protein - huma
331	26	68.4	90	1	VKL0BR	trans-regulatory s
332	26	68.4	91	2	T08665	hypothetical prote
333	26	68.4	94	2	T15982	hypothetical prote
334	26	68.4	97	2	C64022	hypothetical prote
335	26	68.4	97	2	T16463	hypothetical prote
336	26	68.4	100	2	B28487	apolipoprotein C-I
337	26	68.4	108	2	T29281	hypothetical prote
338	26	68.4	113	2	AF1118	hypothetical prote
339	26	68.4	113	2	AH1478	hypothetical prote
340	26	68.4	122	2	AB2199	hypothetical prote
341	26	68.4	124	2	S60044	early nodulin 5 -
342	26	68.4	135	2	F86756	prophage p12 prote
343	26	68.4	137	2	D86768	hypothetical prote
344	26	68.4	139	2	G70464	chlorodioxin - Aqu
345	26	68.4	141	2	D87317	hypothetical prote
346	26	68.4	146	2	B83777	stresses- and starva
347	26	68.4	148	2	S76508	hypothetical prote
348	26	68.4	152	2	A10726	probable membrane
349	26	68.4	152	2	D64943	probable membrane
350	26	68.4	152	2	B99945	hypothetical prote
351	26	68.4	152	2	E85793	hypothetical prote
352	26	68.4	156	2	D75536	ribosomal protein
353	26	68.4	156	2	AE1550	non-heme iron-bind
354	26	68.4	156	2	B82359	DPS family protein
355	26	68.4	156	2	AG1192	non-heme iron-bind
356	26	68.4	157	2	S31634	hypothetical prote
357	26	68.4	157	2	T22831	hypothetical prote
358	26	68.4	158	2	E95036	hypothetical prote
359	26	68.4	173	2	G87383	acetyltransferase,
360	26	68.4	179	2	A11863	hypothetical prote
361	26	68.4	179	2	JU0384	hypothetical 20.2K
362	26	68.4	194	2	T29062	probable transcrip
363	26	68.4	195	2	D82313	2-amino-4-hydroxy-
364	26	68.4	206	2	AB0786	probable membrane
365	26	68.4	207	2	AC2069	acyl carrier prote
366	26	68.4	208	2	AD2213	hypothetical prote
367	26	68.4	212	2	S75535	hypothetical prote
368	26	68.4	214	2	E82679	chaperone XF1452 {
369	26	68.4	214	2	S06455	avenin precursor (
370	26	68.4	224	2	AD0743	probable membrane
371	26	68.4	224	2	C86192	protein T20W3.4 [i
372	26	68.4	227	2	I61843	cytochrome-c-oxida
373	26	68.4	228	2	S75228	hypothetical prote
374	26	68.4	232	2	AG1843	transcription regu
375	26	68.4	232	2	E75559	hypothetical prote
376	26	68.4	232	2	G89025	protein F13A2.8 [i
377	26	68.4	235	2	E95876	conserved hypochet
378	26	68.4	243	2	AF0721	probable Ni/Fe-hyd
379	26	68.4	251	2	PC4261	activin type II re
380	26	68.4	253	2	AE2797	cytochrome C-type
381	26	68.4	254	2	T36161	hypothetical prote
382	26	68.4	255	2	AF3557	transcription regu
383	26	68.4	257	2	E89124	protein K07C11.1 [
384	26	68.4	261	2	T32399	hypothetical prote
385	26	68.4	261	2	F96508	protein F27F5.3 [i
386	26	68.4	265	2	G81241	capsule polysaccha
387	26	68.4	265	2	S15222	ctrc protein - Nei
388	26	68.4	266	2	A86288	protein F9L1.30 [i
389	26	68.4	267	2	E97576	cytochrome C-type
390	26	68.4	272	2	T01144	probable acetone-c
391	26	68.4	278	2	AG2943	hypothetical prote
392	26	68.4	278	2	A83430	transcription regu
393	26	68.4	281	2	T26244	hypothetical prote
394	26	68.4	285	2	C98339	hypothetical prote

395	26	68.4	286	2	D64235
396	26	68.4	287	2	T39197
397	26	68.4	289	2	B69217
398	26	68.4	290	2	T27891
399	26	68.4	292	2	A84935
400	26	68.4	297	2	D95964
401	26	68.4	298	2	T36900
402	26	68.4	298	2	T27365
403	26	68.4	299	2	B95939
404	26	68.4	302	2	AC3020
406	26	68.4	302	2	P98264
407	26	68.4	308	2	T28943
408	26	68.4	309	2	AC3257
409	26	68.4	311	2	T50283
410	26	68.4	314	2	T46743
411	26	68.4	314	2	SG6783
412	26	68.4	315	2	T11388
413	26	68.4	315	2	H86836
414	26	68.4	315	2	T28942
415	26	68.4	317	2	T24346
416	26	68.4	318	2	D69742
417	26	68.4	319	2	B81382
418	26	68.4	319	2	SG4746
419	26	68.4	320	2	T23635
420	26	68.4	320	2	T25308
421	26	68.4	322	2	B84908
422	26	68.4	324	1	A48953
423	26	68.4	324	2	T28940
424	26	68.4	325	2	AC2053
425	26	68.4	326	2	AB3100
426	26	68.4	326	2	H98186
427	26	68.4	328	2	A69425
428	26	68.4	329	2	H83604
429	26	68.4	333	2	T02690
430	26	68.4	338	2	P69035
431	26	68.4	340	2	SS4821
432	26	68.4	341	2	T40424
433	26	68.4	343	2	T40424
434	26	68.4	344	2	B90564
435	26	68.4	344	2	AF3626
436	26	68.4	346	2	A02991
437	26	68.4	346	2	B98292
438	26	68.4	346	2	T24178
439	26	68.4	347	2	A95341
440	26	68.4	349	2	I59336
441	26	68.4	354	2	AF1051
442	26	68.4	355	2	H82481
443	26	68.4	356	2	H84934
444	26	68.4	356	2	H86115
445	26	68.4	356	2	SS6417
446	26	68.4	356	2	H91274
447	26	68.4	357	2	AF2207
448	26	68.4	364	2	UC4249
449	26	68.4	365	2	S27268
450	26	68.4	366	2	A13342
451	26	68.4	368	2	F75546
452	26	68.4	368	2	S18758
453	26	68.4	370	2	T27197
454	26	68.4	375	2	T19169
455	26	68.4	375	2	I48686
456	26	68.4	375	2	A46727
457	26	68.4	378	2	A55735
458	26	68.4	382	2	B49193
459	26	68.4	384	2	T49094
460	26	68.4	395	2	F97023
461	26	68.4	397	2	D72220
462	26	68.4	398	2	T38298
463	26	68.4	400	2	AF0291
464	26	68.4	403	2	C83825
465	26	68.4	406	2	D82968
466	26	68.4	406	2	C75537
467	26	68.4	409	2	A84602

hypothetical prote
yeast atp12 protei
hypothetical prote
hypothetical prote
5,10-methylener
probable transcrip
probable integral
exonzyme S synth
probable spermidin
GDEF family prote
hypothetical prote
integrase/recombin
hypothetical trans
carnitine kinase (c
NADH2 dehydrogen
carnitine kinase (c
hypothetical prote
hypothetical prote
hypothetical prote
tryptophan-tRNA li
CotD8.1 protein -
hypothetical prote
hypothetical prote
cholesterol synth
hypothetical prote
hypothetical prote
probable mikimopin
iron (III) ABC tra
hypothetical prote
hypothetical prote
conserved hypotet
cbrb protein - Brw
hypothetical prote
hypothetical prote
potassium channel
nickel resistance
ABC transporter, m
ribose ABC transpo
hypothetical prote
probable hydrolase
galactin receptor 1
conserved hypotet
conserved hypotet
UDP-N-acetylmutam
hypothetical prote
hypothetical 40.3K
hypothetical prote
hypothetical prote
reep protein - Azo
activin receptor S
D-alanine-D-alanin
conserved hypotet
hypothetical prote
hypothetical prote
N-glycan alpha 2,8
sialyltransferase
G protein-coupled
type II activin re
hypothetical prote
aspartate aminotra
probable aspartate
WD repeat protei
multidrug resistanc
hypothetical prote
conserved hypotet
hypothetical prote

468	26	68.4	412	2	G64059	probable serine tr
469	26	68.4	413	2	E64536	serine transport p
470	26	68.4	417	2	D82172	serine transporter
471	26	68.4	421	2	F83852	glutamate dehydrog
472	26	68.4	422	2	S73369	hypothetical prote
473	26	68.4	428	2	A55044	beta-4C-adrenergic
474	26	68.4	431	2	T24941	hypothetical prote
475	26	68.4	432	2	P96919	ABC transporter (p
476	26	68.4	435	1	S18609	ammonium transport
477	26	68.4	435	2	AC0104	glutamate dehydrog
478	26	68.4	436	2	UC4841	probable sugar tra
479	26	68.4	439	2	G70693	regeneration assoc
480	26	68.4	440	2	T20092	probable dnf prot
481	26	68.4	440	2	G96600	hypothetical prote
482	26	68.4	440	2	G96600	protein F14J16.24
483	26	68.4	443	2	B39794	transcription fact
484	26	68.4	444	2	A39794	transcription fact
485	26	68.4	444	2	B36389	transcription fact
486	26	68.4	444	2	D82328	probable Na+/H+ an
487	26	68.4	445	2	S77010	nickel resistance
488	26	68.4	446	2	T19625	hypothetical prote
489	26	68.4	447	2	T09414	pectinesterase hom
490	26	68.4	450	1	S13730	pmba protein - Esc
491	26	68.4	450	2	F82402	permease VCA0904 [
492	26	68.4	450	2	D86121	maturation of anti
493	26	68.4	450	2	D91280	maturation of anti
494	26	68.4	452	1	G70064	transporter homo
495	26	68.4	457	2	AD3073	hypothetical prote
496	26	68.4	457	2	E98213	hypothetical prote
497	26	68.4	460	2	D72694	hypothetical prote
498	26	68.4	463	2	T39929	thiamin-repressibl
499	26	68.4	469	2	E82264	probable polysacch
500	26	68.4	473	2	AE0801	probable amino aci

ALIGNMENTS

RESULT 1

A56263 beta-galactosidase (EC 3.2.1.23) isozyme 12 - Arthrobacter sp. (strain B7)

C:Species: Arthrobacter sp.

C:Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 16-Aug-2004

C:Accession: A56263

R:Guthrie, K.R.; Trimbur, D.E.; Kasim, J.V.; Brenchley, J.D.

J. Bacteriol. 177, 1981-1988, 1995

A:Title: Analysis of a novel gene and beta-galactosidase isozyme from a psychrotrophic A

A:Reference number: A56263; MUID:95238267; PMID:7721689

A:Accession: A56263

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-637 <GUT>

A:Cross-references: UNIPARC:UPI0000856E1; GB:U17417; NID:ig676845; PID:AAA75601.1; PID:

C:Superfamily: beta-galactosidase, Bgal type

C:Keywords: glycosidase; hydrolase

Query March 89.5%; Score 34; DB 2; Length 637;
Best Local Similarity 85.7%; Pred. No. 32;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 WMLSAFS 7
Db 372 WMLSCFS 378

RESULT 2

T30536 hypothetical protein 2 - Fugu rubripes (fragment)

C:Species: Fugu rubripes

C:Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 09-Jul-2004

C:Accession: T30536

R:Riboldi Tumiciliffe, G.R.; Platzer, M.; Nyakatura, G.; Elgar, G.S.; Brenner, S.; Rosent

submitted to the EMBL Data Library, September 1997

A;Description: Analysis of the genomic loci of Fugu rubripes homologs of the human disea
A;Reference number: Z20848
A;Accession: T30536
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-187 <RTB>
A;Cross-references: UNIPROT:O73698; UNIPARC:UPI000001B8E6; EMBL:AF026198; NID:g3098263;
C;Genetics:
A;Note: PUT2
A;Note: Intron positions not resolved (incomplete sequence)

Query Match 84.2%; Score 32; DB 2; Length 187;
Best Local Similarity 71.4%; Pred. No. 26;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
|||:|:
Db 9 WMLAFA 15

RESULT 3
C70696
probable transport system permease - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: C70696
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.;
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: C70696
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-280 <COL>
A;Cross-references: UNIPROT:P72049; UNIPARC:UPI00000D1237; GB:Z80343; GB:AL123456; NID:9
A;Experimental source: strain H37RV
C;Genetics:
A;Gene: RV3783
C;Superfamily: Integral membrane O-antigen translocator protein rfbA

Query Match 84.2%; Score 32; DB 2; Length 280;
Best Local Similarity 71.4%; Pred. No. 38;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
|||:|:
Db 262 WMLAFA 268

RESULT 4
AB1813
hypothetical protein all0050 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AB1813
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasa moto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimizu, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AB1813
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-345 <KTB>
A;Cross-references: UNIPROT:Q820P0; UNIPARC:UPI00000CDBF9; GB:BA000019; PIDN:BA877574.1;
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all0050

Query Match 84.2%; Score 32; DB 2; Length 345;
Best Local Similarity 71.4%; Pred. No. 46;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
|||:|:
Db 29 WMLSSFS 35

RESULT 5
D83934
hypothetical protein BH2276 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: D83934
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Saeaki, R.; Masui, N.; Fuji, F.; Hirai
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: D83934
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-392 <STO>
A;Cross-references: UNIPROT:Q9KAL1; UNIPARC:UPI00000C3E6F; GB:AP001515; GB:BA000004; NID
C;Genetics:
A;Gene: BH2276

Query Match 84.2%; Score 32; DB 2; Length 392;
Best Local Similarity 71.4%; Pred. No. 53;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
|||:|:
Db 347 WMLQAFS 353

RESULT 6
WZBBB4
51.3K capsid protein - equine herpesvirus 1 (strain AB4p)
C;Species: equine herpesvirus 1
A;Note: host Equus caballus (domestic horse)
C;Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 09-Jul-2004
C;Accession: E36797
R;Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.
submitted to GenBank, March 1992
A;Description: The DNA sequence of equine herpesvirus-1.
A;Reference number: A36805
A;Accession: E36797
A;Molecule type: DNA
A;Residues: 1-465 <TEL>
A;Cross-references: UNIPROT:P28935; UNIPARC:UPI000013B9A; GB:M86664; NID:g330791; PIDN:
R;Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.
Virology 189, 304-316, 1992
A;Title: The DNA sequence of equine herpesvirus-1.
A;Reference number: A41831; MUID:92285566; PMID:1318606
A;Contents: annotation; possible protein-coding frames
A;Note: neither amino acid nor nucleotide sequence is given
C;Genetics:
A;Gene: 22
C;Superfamily: varicella-zoster virus gene 20 protein
C;Keywords: capsid assembly; capsid protein

Query Match 84.2%; Score 32; DB 1; Length 465;
Best Local Similarity 71.4%; Pred. No. 62;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
|||:|:
Db 99 WMLSSFN 105

RESULT 7

D84181
hypothetical protein Vng0208h [imported] - Halobacterium sp. NRC-1
C/Species: Halobacterium sp. NRC-1
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C/Accession: D84181
R/NG: W.V.: Kennedy, S.P.; Mahairas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leitbauer, B.; Keller, K.; Cruz, R.; Danon, M.U.; Hough, D.W.; Maddocks, D.G.; Jaldic
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A/Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A/Title: Genome sequence of Halobacterium species NRC-1.
A/Reference number: A84160; MUID:20504483; PMID:11016950
A/Accession: D84181
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-149 <STD>
A/Cross-references: UNIPROT:Q9H818; UNIPARC:UPI0000063597; GB:AE004437; NID:g10579852; F
C/Genetics:
A/Gene: VNG0208H

Query Match 81.6%; Score 31; DB 2; Length 149;
Best Local Similarity 71.4%; Pred. No. 33;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 WMLSAFS 7
Db 88 WMLDSFS 94

RESULT 8
T51719
hypothetical protein [imported] - Streptococcus mutans
C/Species: Streptococcus mutans
C/Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 09-Jul-2004
C/Accession: T51719
R/Bovd, D.A.; Thervenc, T.; Gumbmann, M.; Honeyman, A.L.; Hamilton, I.R.
Infect. Immun. 68, 925-930, 2000
A/Title: Identification of the operon for the sorbitol (glucitol) phosphoenolpyruvate-su
A/Reference number: 225439; MUID:10107103; PMID:10639465
A/Accession: T51719
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-187 <BOY>
A/Cross-references: UNIPROT:Q9X669; UNIPARC:UPI00000DCB3B; EMBL:AF132127; PIDD:AMD33516.
A/Experimental source: strain LT11

Query Match 81.6%; Score 31; DB 2; Length 187;
Best Local Similarity 83.3%; Pred. No. 41;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WMLSAF 6
Db 10 WMLSAF 15

RESULT 9
H83697
hypothetical protein BH0384 [imported] - Bacillus halodurans (strain C-125)
C/Species: Bacillus halodurans
C/Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C/Accession: H83697
R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hir
Nucleic Acids Res. 28, 4317-4331, 2000
A/Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A/Reference number: A83650; MUID:20512582; PMID:11058132
A/Accession: H83697
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-257 <STD>
A/Cross-references: UNIPROT:Q9KFU2; UNIPARC:UPI00000C386B; GB:AP001508; GB:BA000004; NID
C/Genetics:
A/Experimental source: strain C-125
A/Gene: BH0384

Query Match 81.6%; Score 31; DB 2; Length 257;
Best Local Similarity 83.3%; Pred. No. 56;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WMLSAF 6
Db 183 WMLSAF 188

RESULT 10
T01568
hypothetical protein A_TM018A10.12 - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 09-Jul-2004
C/Accession: T01568
R/Dempsey, S.; Harper, M.
submitted to the EMBL Data Library, July 1997
A/Description: The sequence of A. thaliana TM018A10.
A/Reference number: Z14348
A/Accession: T01568
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-265 <DEM>
A/Cross-references: UNIPROT:Q23098; UNIPARC:UPI00000A37E8; EMBL:AF013294; NID:g2252848; I
C/Genetics:
A/Experimental source: cultivar Columbia
A/Map position: 4
A/Introns: 43/2; 92/1; 186/2
A/Note: AA_TM018A10.12
C/Superfamily: Arabidopsis thaliana hypothetical protein A_TM018A10.12

Query Match 81.6%; Score 31; DB 2; Length 265;
Best Local Similarity 71.4%; Pred. No. 58;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 WMLSAFS 7
Db 160 WMLSKFS 166

RESULT 11
E97674
hypothetical protein AGR_C_4767 [imported] - Agrobacterium tumefaciens (strain C58, Cerec
C/Species: Agrobacterium tumefaciens
C/Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C/Accession: E97674
R/Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Mollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A/Reference number: A97359; MUID:2160851; PMID:11743194
A/Accession: E97674
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-365 <KUR>
A/Cross-references: UNIPROT:Q8U533; UNIPARC:UPI00000D28EF; GB:AE007869; PIDD:AAK88350.1;
C/Genetics:
A/Gene: AGR_C_4767
A/Map position: circular chromosome

Query Match 81.6%; Score 31; DB 2; Length 365;
Best Local Similarity 83.3%; Pred. No. 79;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WMLSAF 6
Db 195 WMLSSF 200

RESULT 12
AD2899
conserved hypothetical protein Atu2629 [imported] - Agrobacterium tumefaciens (strain C5

C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:Accession: AD2899
R:Wood, D.W.; Secubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AD2899
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-368 <KOR>
A:Cross-references: UNIPROT:Q8UC69; UNIPARC:UPI00000D1F84; GB:AE008688; P1DN:AAU43610.1;
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu2629
A:Map position: circular chromosome

Query Match 81.6%; Score 31; DB 2; Length 368;
Best Local Similarity 83.3%; Pred. No. 79;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WMLSAF 6
Db 198 WMLSSF 203

RESULT 13

ORU65
diarylpropane peroxidase (EC 1.11.1.14) H10 precursor - basidiomycete (Phanerochaete chr
N:Alternate names: diarylpropane oxygenase; lignin peroxidase GLG5; ligninase
C:Species: Phanerochaete chrysosporium
C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 12-Jul-2004
C:Accession: JN0117; S13565; E29610; E60995
R:Zhang, Y.Z.; Reddy, C.A.; Raoooly, A.
Gene 97, 191-198, 1991
A:Title: Cloning of several lignin peroxidase (LIP)-encoding genes: sequence analysis of
A:Reference number: JN0117; MUID:91153647; PMID:1999283
A:Accession: JN0117
A:Molecule type: DNA
A:Residues: 1-371 <ZHA>
A:Cross-references: UNIPROT:P11543; UNIPARC:UPI000004CB03; GB:M63496; EMBL:M56815; NID:9
A:Experimental source: strain BKMFI767; ATCC 24725
R:Gaskell, J.; Dieperink, E.; Cullen, D.
Nucleic Acids Res. 19, 599-603, 1991
A:Title: Genomic organization of lignin peroxidase genes of Phanerochaete chrysosporium.
A:Reference number: S13563; MUID:91187681; PMID:12011531
A:Accession: S13565
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-371 <GAS>
A:Cross-references: UNIPARC:UPI000004CB03; GB:X55343; EMBL:X54256; NID:g3137; P1DN:CAA35
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1990
R:de Boer, H.A.; Zhang, Y.Z.; Collins, C.; Reddy, C.A.
Gene 60, 93-102, 1987
A:Title: Analysis of nucleotide sequences of two ligninase cDNAs from a white-rot fungus
A:Reference number: A91587; MUID:8815506; PMID:3440521
A:Accession: B29610
A:Molecule type: mRNA
A:Residues: 1-371 <DEB>
A:Cross-references: UNIPARC:UPI000004CB03; GB:M18743
R:Dass, S.B.; Reddy, C.A.
FEMS Microbiol. Lett. 69, 221-224, 1990
A:Title: Characterization of extracellular peroxidases produced by acetate-buffered cult
A:Reference number: A60995
A:Accession: B60995
A:Molecule type: protein
A:Residues: 28-37 <DAS>
A:Cross-references: UNIPARC:UPI00001721E6
A:Experimental source: ATCC 24725

C:Comment: Lignin peroxidases play a key role in the initial depolymerization and degrad
C:Genetics:
A:Gene: GLG5; LIP6
A:Insertions: 21/3; 71/2; 90/1; 161/1; 175/1; 201/2; 243/2; 343/3; 365/2
C:Superfamily: peroxidase
C:Keywords: extracellular protein; glycoprotein; heme; oxidoreductase
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-27/Domain: propeptide #status predicted <PRO>
F:28-371/Product: lignin peroxidase H10 #status experimental <MAT>
F:283/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 81.6%; Score 31; DB 1; Length 371;
Best Local Similarity 85.7%; Pred. No. 80;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 WMLSAFS 7
Db 197 WMLSAHS 203

RESULT 14

JC1268
lignin peroxidase (EC 1.11.1.-) 2 precursor - basidiomycete (Phanerochaete chrysosporium)
N:Alternate names: lignin peroxidase H6; ligninase H6
C:Species: Phanerochaete chrysosporium
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 12-Jul-2004
C:Accession: JC1268; JH0491; PS0247; C60995
R:Ritch Jr., T.G.; Gold, M.H.
Gene 118, 73-80, 1992
A:Title: Characterization of a highly expressed lignin peroxidase-encoding gene from the
A:Reference number: JC1268; MUID:92380494; PMID:1511887
A:Accession: JC1268
A:Molecule type: DNA
A:Residues: 1-371 <RT>
A:Cross-references: UNIPROT:P49012; UNIPARC:UPI000012B681; GB:M92644; NID:g169277; P1DN:
A:Experimental source: strain OGCI01
R:Ritch Jr., T.G.; Nipper, V.J.; Akileswaran, L.; Smith, A.J.; Pribnow, D.G.; Gold, M.H.
Gene 107, 119-126, 1991
A:Title: Lignin peroxidase from the basidiomycete Phanerochaete chrysosporium is synthe
A:Reference number: JH0491; MUID:92077421; PMID:1743510
A:Accession: JH0491
A:Molecule type: mRNA
A:Residues: 1-371 <RT>
A:Cross-references: UNIPARC:UPI000012B681; GB:M74229; NID:g169271; P1DN:AAA33735.1; P1D:
A:Experimental source: strain OGCI01
A:Accession: PS0247
A:Molecule type: protein
A:Residues: 29-48 <RT1>
A:Cross-references: UNIPARC:UPI0000175214
R:Dass, S.B.; Reddy, C.A.
FEMS Microbiol. Lett. 69, 221-224, 1990
A:Title: Characterization of extracellular peroxidases produced by acetate-buffered cult
A:Reference number: A60995
A:Accession: C60995
A:Molecule type: protein
A:Residues: 28-38 <DAS>
A:Cross-references: UNIPARC:UPI0000175215
A:Experimental source: ATCC 24725
C:Comment: Lignin peroxidase is one of the major components of the lignin degradative sy
C:Genetics:
A:Gene: L62
A:Insertions: 21/1; 72/2; 91/2; 163/1; 177/1; 203/2; 344/3; 366/2
C:Superfamily: peroxidase
C:Keywords: extracellular protein; glycoprotein; heme; oxidoreductase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-28/Domain: propeptide #status predicted <PRO>
F:29-371/Product: lignin peroxidase #status predicted <MAT>
F:295/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 81.6%; Score 31; DB 2; Length 371;
Best Local Similarity 85.7%; Pred. No. 80;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 WMLSAFS 7
|||
Db 199 WMLSAHS 205

RESULT 15

OPUGBP
lignin peroxidase (EC 1.11.1.-) GLG3 precursor - basidiomycete (Phanerochaete chrysosporium)
N/Alternate names: ligninase LIP3
C/Species: Phanerochaete chrysosporium
C/Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 12-Jul-2004
C/Accession: S13723
R/Gaskell, P.S.; Reddy, C.A.
Nucleic Acids Res. 19, 7173, 1990
A/Title: Nucleotide sequence of a new lignin peroxidase gene GLG3 from the white-rot fungus Phanerochaete chrysosporium
A/Reference number: S13723; MUID:91088334; PMID:2129560
A/Accession: S13723
A/Molecule type: DNA
A/Residues: 1-372 <NMI>
A/Cross-references: UNIPROT:P21764; UNIPARC:UPI0000126682; EMBL:X51590; NID:g3135; PIDN: A/Note: the authors translated the codon CAG for residue 5 as Glu
C/Genetics:
A/Gene: GLG3
A/Intons: 21/1; 72/2; 91/2; 163/1; 177/1; 203/2; 344/3; 366/2
C/Superfamily: peroxidase
C/Keywords: extracellular protein; glycoprotein; heme; oxidoreductase
F/1-28/Domain: signal sequence #status predicted <SIG>
F/29-372/Product: lignin peroxidase GLG3 #status predicted <MAT>
F/285/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 81.6%; Score 31; DB 1; Length 372;
Best Local Similarity 85.7%; Pred. No. 80;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 WMLSAFS 7
|||
Db 199 WMLSAHS 205

RESULT 16

OPUGBP
lignin peroxidase (EC 1.11.1.-) LIP3 precursor - basidiomycete (Phanerochaete chrysosporium)
N/Alternate names: ligninase LIP3
C/Species: Phanerochaete chrysosporium
C/Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 12-Jul-2004
C/Accession: S13563
R/Gaskell, J.; Dieperink, E.; Cullen, D.
Nucleic Acids Res. 19, 599-603, 1991
A/Title: Genomic organization of lignin peroxidase genes of Phanerochaete chrysosporium.
A/Reference number: S13563; MUID:91187681; PMID:2011531
A/Accession: S13563
A/Residues: 1-372 <GAS>
A/Molecule type: DNA
A/Status: nucleic acid sequence not shown; translation not shown
A/Cross-references: UNIPROT:P31837; UNIPARC:UPI0000126687; EMBL:X54257; NID:g3154; PIDN: A/Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1990
C/Genetics:
A/Gene: LIP3
A/Intons: 21/1; 72/2; 91/2; 163/1; 177/1; 203/2; 344/3; 366/2
C/Superfamily: peroxidase
C/Keywords: extracellular protein; glycoprotein; heme; oxidoreductase
F/1-28/Domain: signal sequence #status predicted <SIG>
F/29-372/Product: lignin peroxidase LIP3 #status predicted <MAT>
F/285/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 81.6%; Score 31; DB 1; Length 372;
Best Local Similarity 85.7%; Pred. No. 80;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 WMLSAFS 7
|||
Db 199 WMLSAHS 205

RESULT 17

OPUGBP
lignin peroxidase (EC 1.11.1.-) LIPB precursor - basidiomycete (Phanerochaete chrysosporium)
N/Alternate names: ligninase LIP3
C/Species: Phanerochaete chrysosporium
C/Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 12-Jul-2004
C/Accession: S13564
R/Gaskell, J.; Dieperink, E.; Cullen, D.
Nucleic Acids Res. 19, 599-603, 1991
A/Title: Genomic organization of lignin peroxidase genes of Phanerochaete chrysosporium.
A/Reference number: S13563; MUID:91187681; PMID:2011531
A/Accession: S13564
A/Residues: 1-372 <GAS>
A/Cross-references: UNIPROT:P31838; UNIPARC:UPI0000126689; EMBL:X54257; NID:g3154; PIDN: A/Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1990
C/Genetics:
A/Gene: LIPB
A/Intons: 21/1; 72/2; 91/2; 163/1; 177/1; 203/2; 344/3; 366/2
C/Superfamily: peroxidase
C/Keywords: extracellular protein; glycoprotein; heme; oxidoreductase
F/1-28/Domain: signal sequence #status predicted <SIG>
F/29-372/Product: lignin peroxidase LIPB #status predicted <MAT>
F/285/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 81.6%; Score 31; DB 1; Length 372;
Best Local Similarity 85.7%; Pred. No. 80;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 WMLSAFS 7
|||
Db 199 WMLSAHS 205

RESULT 18

S01028
lignin peroxidase (EC 1.11.1.-) LPOA precursor - basidiomycete (Phanerochaete chrysosporium)
N/Alternate names: lignin peroxidase H8; ligninase H8
C/Species: Phanerochaete chrysosporium
C/Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 12-Jul-2004
C/Accession: S01028
R/Smith, T.V.; Schnalch, H.; Gaskell, J.; Covert, S.; Cullen, D.
Nucleic Acids Res. 16, 1219, 1988
A/Title: Nucleotide sequence of a ligninase gene from Phanerochaete chrysosporium.
A/Reference number: S01028; MUID:88144011; PMID:3344218
A/Accession: S01028
A/Molecule type: DNA
A/Residues: 1-372 <SMI>
A/Cross-references: UNIPROT:P06181; UNIPARC:UPI0000166963; EMBL:X06689; NID:g3145; PIDN: A/Note: the authors translated the codon GAC for residue 213 as Asn
C/Genetics:
A/Gene: LPOA
A/Intons: 21/1; 72/2; 91/2; 163/1; 177/1; 203/2; 344/3; 366/2
C/Superfamily: peroxidase
C/Keywords: extracellular protein; glycoprotein; heme; oxidoreductase

Query Match 81.6%; Score 31; DB 2; Length 372;
Best Local Similarity 85.7%; Pred. No. 80;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 WMLSAFS 7
|||
Db 199 WMLSAHS 205

RESULT 19

A36693
lignin peroxidase (EC 1.11.1.-) GLG6 precursor - basidiomycete (Phanerochaete chrysosporium)
C/Species: Phanerochaete chrysosporium
C/Date: 28-Jun-1991 #sequence_revision 28-Jun-1991 #text_change 12-Jul-2004
C/Accession: A36693

C:Accession: A32322; S66556; D60995
R:Schalch, H.; Gaskell, J.; Smith, T.L.; Cullen, D.
Mol. Cell. Biol. 9, 2743-2747, 1989
A:Title: Molecular cloning and sequences of ligin peroxidase genes of Phanerochaete chrysosporium
A:Reference number: A32322; MUID:89343994; PMID:2761543
A:Accession: A32322
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-372 <SCH>
A:Cross-references: UNIPARC:UPI0000175221; GB:M27400
A>Note: the authors translated the codon CAG for residue 110 as His, GAC for residue 191
R:Doyle, W.A.; Smith, A.T.
Biochem. J. 315, 15-19, 1996
A:Title: Expression of ligin peroxidase H8 in Escherichia coli: folding and activation
A:Reference number: S66556; MUID:96207556; PMID:8670100
A:Accession: S66556
A:Molecule type: protein
A:Residues: 22-30 <DOY>
A:Cross-references: UNIPARC:UPI0000175222
R:Dass, S.B.; Reddy, C.A.
FEMS Microbiol. Lett. 69, 221-224, 1990
A:Title: Characterization of extracellular peroxidases produced by acetate-buffered cultures
A:Reference number: A60995
A:Accession: D60995
A:Molecule type: protein
A:Residues: 29-38 <DAS>
A:Cross-references: UNIPARC:UPI0000175223
A:Experimental source: ATCC 24725
A:Superfamily: peroxidase
C:Keywords: extracellular protein; glycoprotein; heme; oxidoreductase

Query Match 81.6%; Score 31; DB 2; Length 372;
Best Local Similarity 85.7%; Pred. No. 80;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 WMLSAFS 7
DB 199 WMLSAHS 205

RESULT 25
J50065
lignin peroxidase (EC 1.11.1.-) 1 precursor - basidiomycete (Phanerochaete chrysosporium)
C:Species: Phanerochaete chrysosporium
C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 12-Jul-2004
C:Accession: J50065
R:Brown, A.; Sims, P.F.G.; Raeder, U.; Broda, P.
Gene 73, 77-85, 1988
A:Title: Multiple liginase-related genes from Phanerochaete chrysosporium.
A:Reference number: J50065; MUID:89211962; PMID:2907500
A:Accession: J50065
A:Molecule type: DNA
A:Residues: 1-372 <BRO>
A:Cross-references: UNIPROT:O01775; UNIPARC:UPI00006A835; GB:M24082; NID:G598354; PIDN:CA7
C:Genetics: LIG1
A:Introns: 21/1; 72/2; 91/2; 163/1; 177/1; 203/2; 344/3; 366/2
C:Superfamily: peroxidase
C:Keywords: extracellular protein; glycoprotein; heme; oxidoreductase
F:1-28/Domain: signal sequence #status predicted <SIG>
F:29-372/Product: ligin peroxidase homolog 1 #status predicted <MAT>

Query Match 81.6%; Score 31; DB 2; Length 372;
Best Local Similarity 85.7%; Pred. No. 80;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 WMLSAFS 7
DB 199 WMLSAHS 205

RESULT 26
B32322

lignin peroxidase (EC 1.11.1.-) 0282 precursor - basidiomycete (Phanerochaete chrysosporium)
C:Species: Phanerochaete chrysosporium
C>Date: 27-Feb-1990 #sequence_revision 27-Feb-1990 #text_change 12-Jul-2004
C:Accession: B32322
R:Schalch, H.; Gaskell, J.; Smith, T.L.; Cullen, D.
Mol. Cell. Biol. 9, 2743-2747, 1989
A:Title: Molecular cloning and sequences of ligin peroxidase genes of Phanerochaete chrysosporium
A:Reference number: A32322; MUID:89343994; PMID:2761543
A:Accession: B32322
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-372 <SCH>
A:Cross-references: UNIPARC:UPI0000175212
A:Superfamily: peroxidase
C:Keywords: extracellular protein; glycoprotein; heme; oxidoreductase

Query Match 81.6%; Score 31; DB 2; Length 372;
Best Local Similarity 85.7%; Pred. No. 80;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 WMLSAFS 7
DB 199 WMLSAHS 205

RESULT 27
JH0156
lignin peroxidase (EC 1.11.1.-) B precursor - basidiomycete (Phanerochaete chrysosporium)
C:Species: Phanerochaete chrysosporium
C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 12-Jul-2004
C:Accession: JH0156
R:Huponen, K.; Ollikka, P.; Kaelin, M.; Walther, I.; Maentzsaale, P.; Reiser, J.
Gene 89, 145-150, 1990
A:Title: Characterization of ligin peroxidase-encoding genes from ligin-degrading basidiomycetes
A:Reference number: JH0156; MUID:90323600; PMID:2373364
A:Accession: JH0156
A:Molecule type: DNA
A:Residues: 1-372 <HUO>
A:Cross-references: UNIPROT:P31837; UNIPARC:UPI0000168967; GB:M37701; GB:M22720; NID:G161
C:Genetics: lpoB
A:Introns: 21/1; 72/2; 91/2; 163/1; 177/1; 203/2; 344/3; 366/2
C:Superfamily: peroxidase
C:Keywords: extracellular protein; glycoprotein; heme; oxidoreductase

Query Match 81.6%; Score 31; DB 2; Length 372;
Best Local Similarity 85.7%; Pred. No. 80;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 WMLSAFS 7
DB 199 WMLSAHS 205

RESULT 28
A27817
lignin peroxidase (EC 1.11.1.-) precursor - basidiomycete (Phanerochaete chrysosporium)
C:Species: Phanerochaete chrysosporium
C>Date: 16-Aug-1988 #sequence_revision 16-Aug-1988 #text_change 12-Jul-2004
C:Accession: A27817
R:Itten, M.; Tu, C.P.D.
Nature 326, 520-523, 1987
A:Title: Cloning and sequencing of a cDNA for a liginase from Phanerochaete chrysosporium
A:Reference number: A27817; MUID:87173020; PMID:3561490
A:Accession: A27817
A:Molecule type: mRNA
A:Residues: 1-573 <TIB>
A:Cross-references: UNIPROT:P06181; UNIPARC:UPI0000168965; GB:Y00262; NID:G3149; PIDN:CA7
C:Superfamily: peroxidase
C:Keywords: extracellular protein; glycoprotein; heme; oxidoreductase
F:1-28/Domain: signal sequence #status predicted <SIG>
F:29-373/Product: ligin peroxidase #status predicted <MAT>

Query Match 81.6%; Score 31; DB 2; Length 373;
Best Local Similarity 85.7%; Pred. No. 80;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
|||
Db 199 WMLSAHS 205

RESULT 29
D96955
sensory transduction histidine kinase (HAMP, HsKA, HATPase domains) [imported] - Clostr
C/Species: Clostridium acetobutylicum
C/Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C/Accession: D96955
R;Nolling, J.; Breton, G.; Omeichenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A/Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cld
A/Reference number: A96900; MUID:21359325; PMID:21359325
A/Accession: D96955
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-416 <KDR>
A/Cross-references: UNIPROT:Q97LV3; UNIPARC:UPI00000C9E8C; GB:AE001437; PIDN:AAK78431.1;
A/Experimental source: Clostridium acetobutylicum ATCC824
A/Genetics:
A/Gene: CAC0451

Query Match 81.6%; Score 31; DB 2; Length 416;
Best Local Similarity 71.4%; Pred. No. 89;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
|||
Db 308 WLEAFS 314

RESULT 30
T48836
hypochemical protein 68B2.200 [imported] - Neurospora crassa
C/Species: Neurospora crassa
C/Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C/Accession: T48836
R;Schulte, U.; Aign, V.; Hohelsel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, April 2000
A/Reference number: Z24541
A/Accession: T48836
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-541 <SCH>
A/Cross-references: UNIPROT:Q9PEV6; UNIPARC:UPI0000069BA5; EMBL:AL353821; GSPDB:GN00112;
A/Experimental source: cosmid contig 68B2, strain 74
C/Genetics:
A/Gene: NCSP:68B2.200
A/Map position: 2
A/Intons: 35/3

Query Match 81.6%; Score 31; DB 2; Length 541;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WMLSAF 6
|||
Db 349 WMLSAF 354

RESULT 31
C82120
transcription regulator Arac/Xy1s family VC2080 [imported] - Vibrio cholerae (strain N16
C/Species: Vibrio cholerae
C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C/Accession: C82120

R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gaitm, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, P
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A/Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A/Reference number: A82035; MUID:20406833; PMID:10952301
A/Accession: C82120
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1121 <HEI>
A/Cross-references: UNIPROT:Q9KQC0; UNIPARC:UPI00000C31D3; GB:AE004282; GB:AE003852; NID
C/Genetics:
A/Experimental source: serogroup O1; strain N16961, biotype El Tor
A/Map position: 1

Query Match 81.6%; Score 31; DB 2; Length 1121;
Best Local Similarity 71.4%; Pred. No. 2.3e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
|||
Db 462 WMLSAFS 468

RESULT 32
B53189
cytochrome b560 homolog precursor - Caenorhabditis elegans
N/Alternate names: hypochemical protein T07C4.7
C/Species: Caenorhabditis elegans
C/Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C/Accession: B53189; S41020
R;Hengartner, M.O.; Horvitz, H.R.
Cell 76, 665-676, 1994
A/Title: Caenorhabditis elegans cell survival gene ced-9 encodes a functional homolog of
A/Reference number: A53189; MUID:94170367; PMID:7907274
A/Accession: B53189
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-182 <HEN>
A/Cross-references: UNIPROT:P41956; UNIPARC:UPI0000126C6F; GB:L26545; NID:9433174; PIDN:
R;Berks, M.
submitted to the EMBL Data Library, January 1994
A/Reference number: S41014
A/Accession: S41020
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-182 <BER>
A/Cross-references: UNIPARC:UPI0000126C6F; EMBL:Z29443; NID:91067051; PID:9443834
C/Genetics:
A/Gene: CYT-1
A/Intons: 29/3; 150/3
C/Superfamily: succinate dehydrogenase cytochrome b560 subunit
C/Keywords: membrane protein; mitochondrion

Query Match 78.9%; Score 30; DB 2; Length 182;
Best Local Similarity 83.3%; Pred. No. 65;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAF 6
|||
Db 67 WMLSGF 72

RESULT 33
T11590
hypochemical protein SPAC17C9.10 - fission yeast (Schizosaccharomyces pombe)
C/Species: Schizosaccharomyces pombe
C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C/Accession: T11590; T52026
R;Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, April 1996
A/Reference number: Z17295

A:Accession: T11590
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-271 <BAR>
A:Cross-references: UNIPROT:Q10482; UNIPARC:UPI0000136110; EMBL:Z73099; NID:g1314152
A:Experimental source: strain 972h(-)
R:Kyung Sook, K.S.
submitted to the EMBL Data Library, January 1999
A:Accession: T52026
A:Reference number: 225907
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-271 <KTU>
A:Cross-references: UNIPARC:UPI0000136110; EMBL:L49134; PIDN:AAD45180.1
A:Experimental source: strain ED65h(-)
C:Genetics:
A:Map position: IR
A>Note: SPAC17C9.10; stml
C:Superfamily: Saccharomyces probable membrane protein YBR147w
C:Keywords: transmembrane protein

Query Match 78.9%; Score 30; DB 2; Length 271;
Best Local Similarity 71.4%; Pred. No. 95;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
Db 246 WILGAFS 252

RESULT 34
D96769
hypothetical protein F9B11.2 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 05-Oct-2004

C:Accession: D96769
R:Theologis, A.; Becker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewart, K.;
ansen, N.F.; Hughes, B.; Hutzar, L.
Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: D96769
A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-305 <STD>

A:Cross-references: UNIPROT:Q9CGA2; UNIPARC:UPI00000A6075; GB:AE005173; NID:g10092421; F

C:Genetics:

A:Gene: F9B11.2

A:Map position: 1

C:Superfamily: AAA protein

Query Match 78.9%; Score 30; DB 2; Length 305;
Best Local Similarity 71.4%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
Db 141 WMLSRYS 147

RESULT 35

G6411
probable tetraacyldisaccharide 4'-kinase (EC 2.7.1.130) HT0059 [similarity] - Haemophilu

C:Species: Haemophilus influenzae

C>Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004

C:Accession: G6411

R:Feischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, M

; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J
D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrman, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995

A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, C

A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.

A:Reference number: A64000; MUID:95350630; PMID:7542800

A:Accession: G6411

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-332 <TTGR>

A:Cross-references: UNIPROT:P44491; UNIPARC:UPI00000378C3; GB:U32691; GB:L42023; NID:g15

A>Note: best homolog was a hypothetical protein from Escherichia coli

C:Superfamily: Escherichia coli probable tetraacyldisaccharide 4'-kinase ycaH

C:Keywords: phosphotransferase

Query Match 78.9%; Score 30; DB 2; Length 332;
Best Local Similarity 71.4%; Pred. No. 1.2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
Db 12 WMLSPFS 18

RESULT 36
C90322
glycosyltransferase [imported] - Sulfolobus solfataricus

C:Species: Sulfolobus solfataricus

C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004

C:Accession: C90322

R:She, Q.; Singh, R.K.; Cafaloni, F.; Zivanovic, Y.; Allard, G.; Aways, M.J.; Chan-Y

Jong, I.; Jeffries, A.C.; Kozera, C.O.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P

arrett, R.A.; Ragan, M.A.; Senses, C.W.; Van der Oost, J.

submitted to GenBank, April 2001

A:Description: Sulfolobus solfataricus complete genome.

A:Reference number: A99139

A:Accession: C90322

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-341 <KUR>

A:Cross-references: UNIPROT:Q97XT6; UNIPARC:UPI00000644FE; GB:AE006641; NID:g13814860; PJ

C:Genetics:

A:Gene: SSO1619

Query Match 78.9%; Score 30; DB 2; Length 341;
Best Local Similarity 71.4%; Pred. No. 1.2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
Db 100 WMLSLFS 106

RESULT 37

B97444
tetraacyldisaccharide 4'-kinase (lipid a 4'-kinase) [imported] - Agrobacterium tumefaciens

C:Species: Agrobacterium tumefaciens

C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004

C:Accession: B97444

R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,

A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.

Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum

A:Reference number: A97359; MUID:2160851; PMID:11743194

A:Accession: B97444

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-348 <KUR>

A:Cross-references: UNIPROT:Q8UH15; UNIPARC:UPI000012E91B; GB:AE007869; PIDN:AAK6507.1;

C:Genetics:

A:Gene: AGR_C1257

A:Map position: circular chromosome

C:Superfamily: Escherichia coli probable tetraacyldisaccharide 4'-kinase ycaH

Query Match 78.9%; Score 30; DB 2; Length 348;
Best Local Similarity 71.4%; Pred. No. 1.2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
|:|:|
|:|:|
Db 18 WMLSPFS 24

RESULT 38
AC2662
tetraacyldisaccharide 4'-kinase [imported] - Agrobacterium tumefaciens (strain C58, Dupc
C/Species: Agrobacterium tumefaciens
C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C/Accession: AC2662
R/Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A/Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A/Reference number: AB2577; MUID:21608550; PMID:11743193
A/Accession: AC2662
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-348 <KUR>
A/Cross-references: UNIPROT:Q8UH15, UNIPARC:UPI000012E91B; GB:AE008688; PIDN:AAL41713.1;
A/Experimental source: strain C58 (Dupont)
C/Genetics:
A/Gene: lpxK
A/Map position: circular chromosome
C/Superfamily: Escherichia coli probable tetraacyldisaccharide 4'-kinase ycaH

Query Match 78.9%; Score 30; DB 2; Length 348;
Best Local Similarity 71.4%; Pred. No. 1.2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
|:|:|
|:|:|
Db 18 WMLSPFS 24

RESULT 39
E82168
probable alpha-1,6-galactosidase VC1690 [imported] - Vibrio cholerae (strain N16961 serc
C/Species: Vibrio cholerae
C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C/Accession: E82168
R/Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, F
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A/Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A/Reference number: A82035; MUID:20406833; PMID:10952301
A/Accession: E82168
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-374 <HEI>
A/Cross-references: UNIPROT:Q9KRF2; UNIPARC:UPI00000C3092; GB:AE004427; GB:AE003852; NID
A/Experimental source: serogroup O1; strain N16961; biotype El Tor
C/Genetics:
A/Gene: VCB1690
A/Map position: 1

Query Match 78.9%; Score 30; DB 2; Length 374;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WMLSAF 6
|:|:|
|:|:|
Db 137 WMLSAF 142

RESULT 40
D82419
tyrosine-specific transport protein VCA0772 [imported] - Vibrio cholerae (strain N16961
C/Species: Vibrio cholerae
C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C/Accession: D82419
R/Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, F
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A/Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A/Reference number: A82035; MUID:20406833; PMID:10952301
A/Accession: D82419
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-400 <HEI>
A/Cross-references: UNIPROT:Q9KLM2; UNIPARC:UPI00000C3641; GB:AE004405; GB:AE003853; NID
A/Experimental source: serogroup O1; strain N16961; biotype El Tor
C/Genetics:
A/Gene: VCA0772
A/Map position: 2
C/Superfamily: tyrosine-specific transport protein

Query Match 78.9%; Score 30; DB 2; Length 400;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
|:|:|
|:|:|
Db 45 WMLMAFT 51

RESULT 41
S65138
glycoprotein antigen MGP57/53, mammary gland - bovine (fragment)
N/Alternate names: glycoprotein component 16/major fat-globule membrane protein/MFG-E8 h
C/Species: Bos primigenius taurus (cattle)
C/Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 07-Aug-1998
C/Accession: S65138; G48394
R/Aoki, N.; Kishi, M.; Taniguchi, Y.; Adachi, T.; Nakamura, R.; Matsuda, T.
Biochim. Biophys. Acta 1245, 385-391, 1995
A/Title: Molecular cloning of glycoprotein antigens MGP57/53 recognized by monoclonal an
A/Reference number: S65138; MUID:96125736; PMID:8541316
A/Accession: S65138
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 1-401 <AOX>
A/Cross-references: UNIPARC:UPI0000177B00
R/Mather, I.H.; Banghart, L.R.; Lane, W.S.
Biochem. Mol. Biol. Int. 29, 545-554, 1993
A/Title: The major fat-globule membrane proteins, bovine components 15/16 and guinea-pig
II-like sequences.
A/Reference number: A48394; MUID:93250576; PMID:8485470
A/Accession: G48394
A/Status: preliminary
A/Molecule type: protein
A/Residues: 207-220 <MAT>
A/Cross-references: UNIPARC:UPI000014C31B
A/Experimental source: milk
A/Note: Sequence extracted from NCBI backbone (NCBI:P:131457)
C/Superfamily: milk fat globule protein; discoidin I amino-terminal homology; EGF homolog
C/Keywords: glycoprotein
F:1-32/Domain: EGF homology (fragment) <EG1>
F:40-79/Domain: EGF homology <EG2>
F:82-239/Domain: discoidin I amino-terminal homology <DNI>
F:243-401/Domain: discoidin I amino-terminal homology <DNI>

Query Match 78.9%; Score 30; DB 2; Length 401;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 WMLSAFS 7
Db 269 WGLSAFS 275

RESULT 42

T11743
pP47 protein - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C:Accession: T11743
R:Enslin, M.; Vogel, T.; Calvete, J.J.; Thole, H.H.; Schmitke, J.; Matsuda, T.; Toesf
Biol. Reprod. 58, 1051-1064, 1998
A:Title: Molecular cloning and characterization of P47, a novel boar sperm-associated z
A:Reference number: Z17325; MUID:98206817; PMID:9546740
A:Accession: T11743
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-409 <ENS>
A:Cross-references: UNIPROT:P79385; UNIPARC:UPI000012F040; EMBL:Y11683; NID:g2652927; P
A:Experimental source: testis
C:Function:
A:Description: may be involved in membrane remodeling and/or function as a zona pellucid
F:6-40/Domain: EGF homology <EGF>

Query Match 78.9%; Score 30; DB 2; Length 409;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 WMLSAFS 7
Db 277 WGLSAFS 283

RESULT 43

F95394
probable transmembrane transport protein SMA937 [imported] - Sinorhizobium meliloti (st
C:Species: Sinorhizobium meliloti
C>Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C:Accession: F95394
R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bowe
; Kaiman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melilot
A:Reference number: A95262; MUID:21396509; PMID:11481432
A:Accession: F95394
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-412 <KUN>
A:Cross-references: UNIPROT:Q92Y26; UNIPARC:UPI00000CB2C5; GB:AEO06469; PIDN:AAK65720.1;
R:Experimental source: strain 1021, megaplasmid pSymA
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebaull, P.; Vandenberg, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A86039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: SMA937
A:Genome: plasmid

Query Match 78.9%; Score 30; DB 2; Length 412;
Best Local Similarity 57.1%; Pred. No. 1.4e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 WMLSAFS 7
Db 53 WGLSAFS 59

RESULT 44

AC0546
probable metabolite transport protein STY0387 [imported] - Salmonella enterica subsp. ent
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A>Note: this species has also been called Salmonella typhi
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AC0546
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
ch, T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skellon, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AC0546
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-417 <PAR>
A:Cross-references: UNIPARC:UPI000005A272; GB:AL513382; PIDN:CAD008811.1; PID:g16501626; C
C:Genetics:
A:Gene: STY0387

Query Match 78.9%; Score 30; DB 2; Length 417;
Best Local Similarity 83.3%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 WMLSAF 6
Db 22 WMLDAF 27

RESULT 45

S74211
PAS-6/7 protein precursor - bovine
N:Alternate names: glycoprotein component 16/major fat-globule membrane protein/MFG-E8 hc
C:Species: Bos primigenius taurus (cattle)
C:Date: 04-Dec-1997 #sequence_revision 12-Dec-1997 #text_change 09-Jul-2004
C:Accession: S74211; S78114; S24181; S65138; G48394
R:Hvattargard, J.; Andersen, M.H.; Berglund, L.; Rasmussen, J.T.; Petersen, T.E.
Eur. J. Biochem. 240, 628-636, 1996
A:Title: Characterization of glycoprotein PAS-6/7 from membranes of bovine milk fat globu
A:Reference number: S74211; MUID:97008954; PMID:8856064
A:Accession: S74211
A:Molecule type: mRNA
A:Residues: 1-427 <HVA>
A:Cross-references: UNIPROT:Q95114; UNIPARC:UPI000016C359; EMBL:X91895; NID:g1632778; PII
A:Accession: S78114
A:Molecule type: protein
A:Residues: 19-85; 96-110; 140-165; 174-216; 221-232; 248-277; 285-293; 309-337; 339-420; 425-427
A:Cross-references: UNIPARC:UPI0000177B01; UNIPARC:UPI0000177B02; UNIPARC:UPI0000177B03;
B08; UNIPARC:UPI0000177B09; UNIPARC:UPI0000177B0A
R:Kim, D.H.; Kanno, C.; Mizokami, Y.
Biochim. Biophys. Acta 1122, 203-211, 1992
A:Title: Purification and characterization of major glycoproteins, PAS-6 and PAS-7, from
A:Reference number: S23926; MUID:92353107; PMID:1643094
A:Accession: S24181
A:Molecule type: protein
A:Residues: 383-394 <KIM>
A:Cross-references: UNIPARC:UPI0000177B0B
R:Aoki, N.; Kishi, M.; Taniguchi, Y.; Adachi, T.; Nakamura, R.; Matsuda, T.
Biochim. Biophys. Acta 1245, 385-391, 1995
A:Title: Molecular cloning of glycoprotein antigens MGP57/53 recognized by monoclonal ant
A:Reference number: S65138; MUID:96125736; PMID:8541316
A:Accession: S65138
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 27-427 <AOK>
A:Cross-references: UNIPARC:UPI0000177B00
R:Maicher, I.H.; Banghart, L.R.; Lane, W.S.
Biochem. Mol. Biol. Int. 29, 545-554, 1993
A:Title: The major fat-globule membrane proteins, bovine components 15/16 and guinea-pig
II-like sequences.

A:Reference number: A48394; MUID:33250576; PMID:8485470
A:Accession: G48394
A:Status: preliminary
A:Molecule type: protein
A:Residues: 233-246 <MAT>
A:Cross-references: UNIPARC:UPI000014C318
A:Experimental source: milk
A>Note: sequence extracted from NCBI backbone (NCBIP:131457)
C:Superfamily: milk fat globule protein; discoidin I amino-terminal homology; EGF homolo
C:Keywords: blocked amino end; disulfide bond; glycoprotein; milk
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-427/Product: PAS-6/7 protein #status experimental <MAT>
F:24-58/Domain: EGF homology <EG1>
F:66-105/Domain: EGF homology <EG2>
F:108-265/Domain: discoidin I amino-terminal homology <DN1>
F:269-427/Domain: discoidin I amino-terminal homology <DN2>
F:24-35-29-47-49-58-66-77-71-94-96-105/Disulfide bonds: #status predicted
F:27/Binding site: carbohydrate (Ser) (covalent) #status experimental
F:34/Binding site: carbohydrate (Thr) (covalent) #status experimental
F:59-227/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:109-265-252-256-270-427/Disulfide bonds: #status experimental

Query Match 78.9%; Score 30; DB 2; Length 427;
Best Local Similarity 85.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 WMLSAFS 7
Db 295 WGLSAFS 301

RESULT 46
A75444
hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: A75444
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathavan, J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Frazer, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: A75444
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-454 <WHI>
A:Cross-references: UNIPROT:Q9RVH3; UNIPARC:UPI0000C187D; GB:AE001956; GB:AE000513; NID
C:Experimental source: strain R1
C:Genetics:
A:Gene: DR1056
A:Map position: 1
C:Superfamily: citrate utilization determinant

Query Match 78.9%; Score 30; DB 2; Length 454;
Best Local Similarity 83.3%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WMLSAF 6
Db 54 WMLSAF 59

RESULT 47
F90544
conserved hypothetical protein MYPV_2620 [imported] - Mycoplasma pulmonis (strain UAB CT
C:Species: Mycoplasma pulmonis
C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C:Accession: F90544
R:Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.;
Nucleic Acids Res. 29, 2145-2153, 2001
A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pult
A:Reference number: A99512; MUID:21267165; PMID:11353084

A:Accession: F90544
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-467 <KUR>
A:Cross-references: UNIPROT:Q96Q09; UNIPARC:UPI0000C8046; GB:AL445566; PID:q14089675; PJ
A:Experimental source: strain UAB CTIP
C:Genetics:
A:Gene: MYPV_2620
A:Genetic code: SGC3
C:Superfamily: hypothetical protein MG294

Query Match 78.9%; Score 30; DB 2; Length 467;
Best Local Similarity 83.3%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 WMLSAF 6
Db 246 WMLSVF 251

RESULT 48
S70171
integral membrane protein - Streptomyces pristinaespiralis
C:Species: Streptomyces pristinaespiralis
C>Date: 19-Mar-1997 #sequence_revision 01-Aug-1997 #text_change 31-Dec-2004
C:Accession: S70171; S57509
R:Blanc, V.; Salah-Bey, K.; Folcher, M.; Thompson, C.J.
Mol. Microbiol. 17, 989-999, 1995
A:Title: Molecular characterization and transcriptional analysis of a multidrug resistant
A:Reference number: S70171; MUID:96123441; PMID:8596448
A:Accession: S70171
A:Molecule type: DNA
A:Residues: 1-501 <BLA>
A:Cross-references: UNIPROT:Q54806; UNIPARC:UPI0000B6054; EMBL:X84072; NID:9872305; PID
C:Genetics:
A:Gene: ptr

Query Match 78.9%; Score 30; DB 2; Length 501;
Best Local Similarity 57.1%; Pred. No. 1.7e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 WMLSAFS 7
Db 60 WMLNAYS 66

RESULT 49
T40598
hypothetical protein SPBC649.05 - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T40598
R:Purnelle, B.; Goffeau, A.; Lyne, M.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, May 1998
A:Reference number: Z21939
A:Accession: T40598
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-548 <PUR>
A:Cross-references: UNIPROT:O59755; UNIPARC:UPI0000128686; EMBL:AL023587; PIDN:CAA19047.1
A:Experimental source: strain 972h-; cosmid c649
C:Genetics:
A:Gene: SPBC:SPBC649.05
A:Map position: 2
A:introns: 3/1

Query Match 78.9%; Score 30; DB 2; Length 548;
Best Local Similarity 71.4%; Pred. No. 1.9e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 WMLSAFS 7
Db 13 WMLKAFS 19

RESULT 50

A39280
glutamate receptor GluR-II precursor - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 28-Feb-1992 #sequence_revision 10-Apr-1992 #text_change 31-Dec-2004
C;Accession: A39280
R;Schuster, C.M.; Ultsch, A.; Schloess, P.; Cox, J.A.; Schmitt, B.; Betz, H.
Science 254, 112-114, 1991
A;Title: Molecular cloning of an invertebrate glutamate receptor subunit expressed in D.
A;Reference number: A39280; MUID:92022535; PMID:1681587
A;Accession: A39280
A;Status: Preliminary
A;Molecule type: mRNA
A;Residues: 1-906 <SCH>
A;Cross-references: UNIPROT:Q24351; UNIPROT:Q9VMP4; UNIPARC:UPI0000177962; GB:M73271
C;Genetics:
A;Gene: FlyBase:Glu-RII
A;Cross-references: FlyBase:FBgn0004620
C;Superfamily: mannose 6-phosphate receptor, cation-independent; glutamate receptor homodimer
C;Keywords: neurotransmitter receptor
F;431-852/Domain: glutamate receptor homology <GRH>

Query Match 78.9%; Score 30; DB 2; Length 906;
Best local Similarity 83.3%; Pred.No. 3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 WMLSAF 6
Db 344 WMLGAF 349

Search completed: August 29, 2006, 06:14:15
Job time : 17 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2006, 05:47:27 ; Search time 56 Seconds
(without alignments)
115.627 Million cell updates/sec

Title: US-10-541-343-3

Perfect score: 38

Sequence: 1 WMLSAPS 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 500 summaries

Database :

UniProt 7.2:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	92.1	71	2	084G66_CORYGU
2	35	92.1	347	2	08N0Z0_HUMAN
3	35	92.1	424	2	05L241_GEOXA
4	35	92.1	459	2	09B6P3_THRSW
5	35	92.1	1760	2	09VLT3_DROME
6	35	92.1	2943	2	034A91_RHOPA
7	34	89.5	147	2	04MR97_BACCE
8	34	89.5	147	2	063CE7_BACCE
9	34	89.5	147	2	06HJV3_BACHK
10	34	89.5	147	2	0813P2_BACCR
11	34	89.5	147	2	0739P8_BACCI
12	34	89.5	147	2	08RPO1_BACCI
13	34	89.5	154	2	03EX04_BACCI
14	34	89.5	211	2	03ES54_VIBFI
15	34	89.5	361	2	0744H1_MYCPA
16	34	89.5	458	2	02W9C2_MAGSA
17	34	89.5	637	2	044233_ARTSP
18	34	89.5	843	2	088944_HHRV
19	34	89.5	853	2	03WVR2_PACCN
20	34	89.5	1030	2	073RB9_TREDE
21	34	89.5	1196	2	04B833_BURVI
22	34	86.8	44	2	05QTR8_HUMAN
23	33	86.8	304	2	02UD08_ASPOR
24	33	86.8	1404	1	PRG4_HUMAN
25	32	84.2	74	2	046TT1_PALEJ
26	32	84.2	117	2	04S8Z0_TESTG
27	32	84.2	129	2	043Q96_SOLUS
28	32	84.2	159	2	07MXJ3_PORGI
29	32	84.2	159	2	07MXS1_PORGI
30	32	84.2	167	2	084S36_ORYSA
31	32	84.2	182	2	04SFB2_TESTG

32	32	84.2	187	1	PUR2_FUGRU	073698 fungu rubrip
33	32	84.2	226	2	06CXD3_KULUA	06CXD3 Kluveromyc
34	32	84.2	243	2	06ZRV1_HUMAN	06ZRV1 homo sapien
35	32	84.2	280	2	P72049_MYCTU	P72049 mycobacteri
36	32	84.2	280	2	07TVN9_MYCBO	07TVN9 mycobacteri
37	32	84.2	289	1	MRGRG_MOUSE	MRGRG mouse
38	32	84.2	289	1	0498A2_MOUSE	0498A2 mus musculu
39	32	84.2	333	2	03M9S1_ANAVT	03M9S1 anabena va
40	32	84.2	341	2	02TNR7_9DIPT	02TNR7 dirosophila
41	32	84.2	341	2	02TNS2_9DIPT	02TNS2 dirosophila
42	32	84.2	341	2	02TNS6_9DIPT	02TNS6 dirosophila
43	32	84.2	341	2	02TNT2_9DIPT	02TNT2 dirosophila
44	32	84.2	341	2	02TNT3_9DIPT	02TNT3 dirosophila
45	32	84.2	341	2	02TNT5_9DIPT	02TNT5 dirosophila
46	32	84.2	341	2	02TNT6_9DIPT	02TNT6 dirosophila
47	32	84.2	341	2	02TNT7_DROKX	02TNT7 dirosophila
48	32	84.2	341	2	02TNT9_9DIPT	02TNT9 dirosophila
49	32	84.2	341	2	02TNU0_9DIPT	02TNU0 dirosophila
50	32	84.2	341	2	02TNU4_9DIPT	02TNU4 dirosophila
51	32	84.2	341	2	035CL6_9BRAD	035CL6 bradyrhizob
52	32	84.2	341	2	08Z0P0_ANASP	08Z0P0 anabena sp
53	32	84.2	342	2	09KAL1_BACHD	09KAL1 bacillus ha
54	32	84.2	410	2	07MMW4_BORBR	07MMW4 bordetella
55	32	84.2	416	2	07WBE3_BORPA	07WBE3 bordetella
56	32	84.2	452	2	07VUX6_BORPE	07VUX6 bordetella
57	32	84.2	465	1	VP19_EHY1B	VP19 equine herp
58	32	84.2	465	1	VP19_EHY1V	VP19 equine herp
59	32	84.2	465	1	06Z4A6_ORYSA	06Z4A6 oryza sativ
60	32	84.2	541	2	06C2L1_YARLI	06C2L1 yarrowia li
61	32	84.2	551	2	04BH19_BURVI	04BH19 burkholderi
62	32	84.2	564	2	06CB15_YARLI	06CB15 yarrowia li
63	32	84.2	569	2	062DU9_BURMA	062DU9 burkholderi
64	32	84.2	597	2	063UB8_BURPS	063UB8 burkholderi
65	32	84.2	597	2	07WTD1_PACYO	07WTD1 streptomyc
66	32	84.2	611	2	03JK96_BURPI	03JK96 burkholderi
67	32	84.2	740	2	06K613_ORYSA	06K613 oryza sativ
68	32	84.2	747	2	06EUL7_ORYSA	06EUL7 oryza sativ
69	32	84.2	782	2	07SE61_NEUCR	07SE61 neurospora
70	32	84.2	872	2	02URH0_ASPOR	02URH0 aspergillus
71	32	84.2	876	2	03F1S6_9BURK	03F1S6 burkholderi
72	32	84.2	1196	2	044U11_9BURK	044U11 burkholderi
73	32	84.2	1196	2	04LIA7_9BURK	04LIA7 burkholderi
74	32	84.2	1196	2	039KJ6_BURR3	039KJ6 burkholderi
75	32	84.2	1595	2	02USK9_ASPOR	02USK9 aspergillus
76	32	84.2	1595	2	092219_CERSU	092219 cerioporios
77	31	81.6	68	2	092220_CERSU	092220 cerioporios
78	31	81.6	68	2	092267_PAPHY	092267 planerococcae
79	31	81.6	70	2	03RUQ2_RALME	03RUQ2 ralsconia m
80	31	81.6	86	2	04AP45_GCHLB	04AP45 chlorobium
81	31	81.6	89	2	03XBG0_METFL	03XBG0 methylobacti
82	31	81.6	89	2	02IEB4_9DELT	02IEB4 anaeromyxob
83	31	81.6	99	2	09SPH2_GOSHI	09SPH2 gossypium h
84	31	81.6	143	2	09HS18_HALSA	09HS18 halobacteri
85	31	81.6	149	2	071172_LACDI	071172 lactobacilli
86	31	81.6	151	2	05BZT8_SCHJA	05BZT8 schistosoma
87	31	81.6	152	2	06SF57_BACLD	06SF57 bacillus li
88	31	81.6	154	2	06E4A5_CYNDA	06E4A5 cynodon dac
89	31	81.6	171	2	082RP6_STRAW	082RP6 streptomyc
90	31	81.6	178	2	044G05_CHRSL	044G05 chromohalob
91	31	81.6	181	2	09X669_STRMU	09X669 streptococc
92	31	81.6	187	2	087IS4_VIBPA	087IS4 vibrio para
93	31	81.6	210	2	03CVR7_ALTRA	03CVR7 pseudalter
94	31	81.6	214	2	0425Y0_PLABE	0425Y0 plasmodium
95	31	81.6	214	2	08NKH4_PAGAR	08NKH4 tricholoma
96	31	81.6	249	2	09KFU2_BACHD	09KFU2 bacillus ha
97	31	81.6	253	2	06GPI1_HUMAN	06GPI1 homo sapien
98	31	81.6	263	2	023098_ARATH	023098 arabidopsis
99	31	81.6	265	2	09RCJ3_BACHD	09RCJ3 bacillus ha
100	31	81.6	274	2	08LAS4_ARATH	08LAS4 arabidopsis
101	31	81.6	290	2	04NHS6_PMITC	04NHS6 atrobacte
102	31	81.6	299	2	08VZ38_ARATH	08VZ38 arabidopsis
103	31	81.6	302	2	07XAM5_ORYSA	07XAM5 oryza sativ
104	31	81.6	302	2		

105	31	81.6	317	2	Q5XVF8_ARATH	Q5XVF8_arabidopsis	178	30	78.9	141	2	Q3YK21_CVHNL	Q3YK21_human
106	31	81.6	328	2	Q6ZJ45_BURMA	Q6ZJ45_burkholderi	179	30	78.9	141	2	Q3YK22_CVHNL	Q3YK22_human
107	31	81.6	342	2	Q4FKX6_PELUB	Q4FKX6_pelagibacte	180	30	78.9	141	2	Q3YK23_CVHNL	Q3YK23_human
108	31	81.6	345	2	Q7LHY5_PHACH	Q7LHY5_phanerochae	181	30	78.9	141	2	Q3YK24_CVHNL	Q3YK24_human
109	31	81.6	347	2	Q43MT9_SOLUS	Q43MT9_solibacter	182	30	78.9	141	2	Q3YK25_CVHNL	Q3YK25_human
110	31	81.6	347	2	Q4QYG7_98AUR	Q4QYG7_pygopus lep	183	30	78.9	141	2	Q3YK26_CVHNL	Q3YK26_human
111	31	81.6	362	2	Q01785_PHACH	Q01785_phanerochae	184	30	78.9	141	2	Q3YK27_CVHNL	Q3YK27_human
112	31	81.6	365	2	Q8U533_AGR75	Q8U533_agrobacteri	185	30	78.9	141	2	Q3YK28_CVHNL	Q3YK28_human
113	31	81.6	368	2	Q8UC69_AGR75	Q8UC69_agrobacteri	186	30	78.9	141	2	Q3YK29_CVHNL	Q3YK29_human
114	31	81.6	371	1	LIG2_PHACH	P49011_phanerochae	187	30	78.9	141	2	Q3YK30_CVHNL	Q3YK30_human
115	31	81.6	371	1	LIG5_PHACH	P11543_phanerochae	188	30	78.9	141	2	Q3YK31_CVHNL	Q3YK31_human
116	31	81.6	371	2	Q01787_PHACH	Q01787_phanerochae	189	30	78.9	141	2	Q3YK32_CVHNL	Q3YK32_human
117	31	81.6	371	2	Q06326_PHACH	Q06326_phanerochae	190	30	78.9	141	2	Q3YK33_CVHNL	Q3YK33_human
118	31	81.6	372	1	LIG3_PHACH	P21764_phanerochae	191	30	78.9	141	2	Q5XNK6_CVHNL	Q5XNK6_human
119	31	81.6	372	1	LIG6_PHACH	P50622_phanerochae	192	30	78.9	141	2	Q5XNK7_CVHNL	Q5XNK7_human
120	31	81.6	372	1	LIG8_PHACH	P06181_phanerochae	193	30	78.9	141	2	Q5XNK8_CVHNL	Q5XNK8_human
121	31	81.6	372	1	LIGA_PHACH	P31837_phanerochae	194	30	78.9	141	2	Q5XNK9_CVHNL	Q5XNK9_human
122	31	81.6	372	1	LIGB_PHACH	P31838_phanerochae	195	30	78.9	141	2	Q5XNL0_CVHNL	Q5XNL0_human
123	31	81.6	372	2	Q01775_PHACH	Q01775_phanerochae	196	30	78.9	141	2	Q5XNL1_CVHNL	Q5XNL1_human
124	31	81.6	392	2	Q40G54_98HOB	Q40G54_jamaichia	197	30	78.9	141	2	Q5XNL2_CVHNL	Q5XNL2_human
125	31	81.6	406	2	Q63ST5_BURPS	Q63ST5_burkholderi	198	30	78.9	141	2	Q5XNL3_CVHNL	Q5XNL3_human
126	31	81.6	416	2	Q3CG27_THBET	Q3CG27_thermoan aer	199	30	78.9	141	2	Q6B7R4_CVHNL	Q6B7R4_human
127	31	81.6	416	2	Q97LV3_CLOAB	Q97LV3_clostridium	200	30	78.9	141	2	Q6B7R5_CVHNL	Q6B7R5_human
128	31	81.6	421	2	Q5OSM7_ENTHT	Q5OSM7_entamoeba h	201	30	78.9	141	2	Q6B7R6_CVHNL	Q6B7R6_human
129	31	81.6	430	2	Q3JQU4_BURPI	Q3JQU4_burkholderi	202	30	78.9	141	2	Q6B7R7_CVHNL	Q6B7R7_human
130	31	81.6	436	2	Q3P6P2_9GAMM	Q3P6P2_shewanella	203	30	78.9	141	2	Q6B7R8_CVHNL	Q6B7R8_human
131	31	81.6	440	2	Q5LJY0_DROME	Q5LJY0_drosophila	204	30	78.9	141	2	Q6B7R9_CVHNL	Q6B7R9_human
132	31	81.6	441	2	Q7PLM4_DROME	Q7PLM4_drosophila	205	30	78.9	141	2	Q6B7S0_CVHNL	Q6B7S0_human
133	31	81.6	444	2	Q8OPF5_RAJEG	Q8OPF5_raja eglant	206	30	78.9	141	2	Q6B7S1_CVHNL	Q6B7S1_human
134	31	81.6	456	2	Q2RP48_RHORU	Q2RP48_rhodospirill	207	30	78.9	141	2	Q6B7S2_CVHNL	Q6B7S2_human
135	31	81.6	458	2	Q7RL69_PLAYO	Q7RL69_plasmodium	208	30	78.9	141	2	Q6B7S3_CVHNL	Q6B7S3_human
136	31	81.6	472	2	Q8BX72_LACPL	Q8BX72_lactobacill	209	30	78.9	141	2	Q6B7S4_CVHNL	Q6B7S4_human
137	31	81.6	494	2	Q7NNF4_GLOVI	Q7NNF4_gloeobacter	210	30	78.9	141	2	Q6B7S5_CVHNL	Q6B7S5_human
138	31	81.6	494	2	Q4V9P5_BRABE	Q4V9P5_brachydantio	211	30	78.9	141	2	Q6B7S6_CVHNL	Q6B7S6_human
139	31	81.6	541	2	Q9P6V6_TOXCO	Q9P6V6_neurospora	212	30	78.9	149	2	P83695_BREBE	P83695_brevibacill
140	31	81.6	558	2	Q6BVO5_TOXCO	Q6BVO5_toxoplasma	213	30	78.9	150	2	Q3CTX2_ALTTAT	Q3CTX2_pseudotaler
141	31	81.6	640	2	Q5P9M9_ANAMM	Q5P9M9_anaplasm a	214	30	78.9	155	2	Q6Q1R1_CVHNL	Q6Q1R1_human
142	31	81.6	651	2	Q4S4S0_TETNG	Q4S4S0_tetradon n	215	30	78.9	155	2	Q6Q1R2_CVHNL	Q6Q1R2_human
143	31	81.6	802	2	Q8CEX0_MOUSE	Q8CEX0_mus musculu	216	30	78.9	155	2	Q6Q1R3_CVHNL	Q6Q1R3_human
144	31	81.6	848	2	Q2NFH1_9EURY	Q2NFH1_mechanospha	217	30	78.9	155	2	Q6Q1R4_CVHNL	Q6Q1R4_human
145	31	81.6	999	2	Q4HZU8_GIBZE	Q4HZU8_glibberella	218	30	78.9	155	2	Q6Q1R5_CVHNL	Q6Q1R5_human
146	31	81.6	1121	2	Q9KOCO_VIBCH	Q9KOCO_vibrio chol	219	30	78.9	155	2	Q6Q1R6_CVHNL	Q6Q1R6_human
147	31	81.6	1141	2	Q4QBL6_LEIMA	Q4QBL6_leishmania	220	30	78.9	155	2	Q6Q1R7_CVHNL	Q6Q1R7_human
148	31	81.6	1157	2	Q3WU51_9RHIZ	Q3WU51_mesorhizobi	221	30	78.9	159	2	Q77553_PIG	Q77553_sus scrofa
149	31	81.6	1183	2	Q5UUG2_CRYNE	Q5UUG2_cryptococcu	222	30	78.9	167	2	Q63V54_BURPS	Q63V54_burkholderi
150	31	81.6	1183	2	Q5KHY2_CRYNE	Q5KHY2_cryptococcu	223	30	78.9	169	2	Q3QCBO_9GAMM	Q3QCBO_shewanella
151	31	81.6	1464	2	Q4HZM0_GIBZE	Q4HZM0_glibberella	224	30	78.9	169	2	Q637R2_BURPS	Q637R2_burkholderi
152	31	81.6	1687	2	Q514M1_ENTHT	Q514M1_entamoeba h	225	30	78.9	171	2	Q4HA50_TDEIO	Q4HA50_deinococcus
153	30	78.9	79	2	Q3SWM6_MARHY	Q3SWM6_marinobacte	226	30	78.9	172	2	Q3B1W0_PBLD	Q3B1W0_peloidictyon
154	30	78.9	101	2	Q7UVT0_RHOBA	Q7UVT0_rhodospirill	227	30	78.9	176	2	Q9TV80_PIG	Q9TV80_sus scrofa
155	30	78.9	112	2	Q4HLM4_CAMLA	Q4HLM4_campylobact	228	30	78.9	177	2	Q77546_PIG	Q77546_sus scrofa
156	30	78.9	116	2	Q84UM9_CERHO	Q84UM9_ceratiu ho	229	30	78.9	182	1	C560_CABEL	P41956_caenothabdi
157	30	78.9	116	2	Q3OBX5_9H1V1	Q3OBX5_human immun	230	30	78.9	182	2	Q5VL77_LEIDO	Q5VL77_leishmania
158	30	78.9	116	2	Q91BN5_9H1V1	Q91BN5_human immun	231	30	78.9	182	2	Q61MM3_LEIMA	Q61MM3_leishmania
159	30	78.9	116	2	Q2MMQ4_9H1V1	Q2MMQ4_human immun	232	30	78.9	184	1	C560_CABER	P41955_caenothabdi
160	30	78.9	126	2	Q3P225_9GAMM	Q3P225_shewanella	233	30	78.9	237	2	Q2KMP1_CVHNL	Q2KMP1_human
161	30	78.9	127	2	Q3EUD0_BACTI	Q3EUD0_bacillus th	234	30	78.9	238	2	Q5DIY8_CVHNL	Q5DIY8_human
162	30	78.9	127	2	Q5L227_GROKA	Q5L227_geobacillus	235	30	78.9	238	2	Q5DIY9_CVHNL	Q5DIY9_human
163	30	78.9	127	2	Q6RCH1_BACHK	Q6RCH1_bacillus th	236	30	78.9	238	2	Q5DIZ0_CVHNL	Q5DIZ0_human
164	30	78.9	127	2	Q816Q6_BACCR	Q816Q6_bacillus ce	237	30	78.9	238	2	Q5DIZ1_CVHNL	Q5DIZ1_human
165	30	78.9	127	2	Q81KH6_BACAN	Q81KH6_bacillus ac	238	30	78.9	238	2	Q5DIZ3_CVHNL	Q5DIZ3_human
166	30	78.9	131	2	Q3MYI8_9DELT	Q3MYI8_syntrophoba	239	30	78.9	238	2	Q5DIZ4_CVHNL	Q5DIZ4_human
167	30	78.9	138	2	Q3UG79_MOUSE	Q3UG79_mus musculu	240	30	78.9	240	2	Q36JG5_MARHY	Q36JG5_marinobacte
168	30	78.9	139	2	Q9CZU9_MOUSE	Q9CZU9_mus musculu	241	30	78.9	244	2	Q5X3B1_LEGPA	Q5X3B1_gloeobacter
169	30	78.9	141	2	Q3OHR3_CVHNL	Q3OHR3_human	242	30	78.9	247	2	Q7NIR3_GLOVI	Q7NIR3_gloeobacter
170	30	78.9	141	2	Q3OHR4_CVHNL	Q3OHR4_human	243	30	78.9	251	2	Q211A2_CVHNL	Q211A2_human
171	30	78.9	141	2	Q3OHR5_CVHNL	Q3OHR5_human	244	30	78.9	251	2	Q211A3_CVHNL	Q211A3_human
172	30	78.9	141	2	Q3OHR6_CVHNL	Q3OHR6_human	245	30	78.9	260	2	Q3UW20_MOUSE	Q3UW20_mus musculu
173	30	78.9	141	2	Q3OHR7_CVHNL	Q3OHR7_human	246	30	78.9	260	2	Q3UW29_MOUSE	Q3UW29_mus musculu
174	30	78.9	141	2	Q3OHR8_CVHNL	Q3OHR8_human	247	30	78.9	262	2	Q81OU1_MOUSE	Q81OU1_mus musculu
175	30	78.9	141	2	Q3OHR9_CVHNL	Q3OHR9_human	248	30	78.9	267	1	Q947B8_ORYSA	Q947B8_oryza sativ
176	30	78.9	141	2	Q3OHS0_CVHNL	Q3OHS0_human	249	30	78.9	271	1	STM1_SCHPO	STM1_schizosacch
177	30	78.9	141	2	Q3YK20_CVHNL	Q3YK20_human	250	30	78.9	271	2	Q5XKZ8_AERHY	Q5XKZ8_aetomonas h

251	30	78.9	276	2	060M14_CAEBR	060M14_caenorhabdi	324	30	78.9	1128	2	060EX7_ORYSA	060EX7_oryza sativ
252	30	78.9	283	2	051M15_GILPO	051M15_gillichbcte	325	30	78.9	1161	2	048J17_PSE14	048J17_pseudomonas
253	30	78.9	288	2	03VB55_9SPHN	03VB55_sphingopyxi	326	30	78.9	1190	2	04V0A7_XANC8	04V0A7_xanthomonas
254	30	78.9	298	2	098HR3_RH110	098HR3_rhizobium 1	327	30	78.9	1190	2	08PE13_XANCP	08PE13_xanthomonas
255	30	78.9	302	2	051HB6_BACFN	051HB6_bacteroides	328	30	78.9	1321	2	04JFH3_CAEEL	04JFH3_caenorhabdi
256	30	78.9	302	2	064Y85_BACFN	064Y85_bacteroides	329	30	78.9	1350	2	061YK0_CAEEL	061YK0_caenorhabdi
257	30	78.9	305	2	09C6A2_ARATH	09C6A2_arabidopsis	330	30	78.9	1372	2	020320_CAEEL	020320_caenorhabdi
258	30	78.9	313	2	03JTH5_BURP1	03JTH5_burkholderi	331	30	78.9	1525	2	070YF3_GIALA	070YF3_giardia lam
259	30	78.9	317	2	07S102_NEUCR	07S102_neurospora	332	30	78.9	1660	2	06C8Z8_YARL1	06C8Z8_yarrowia li
260	30	78.9	318	2	02K576_RH1ET	02K576_rhizobium e	333	30	78.9	1809	2	08T398_C10IN	08T398_clona inres
261	30	78.9	324	2	08HQJ2_ORNMO	08HQJ2_ornithodoros	334	30	78.9	2140	2	05CWC2_CRYPY	05CWC2_cryptospori
262	30	78.9	330	2	07YXL2_TENMO	07YXL2_tenebrio mo	335	30	78.9	2142	2	05CHM9_CRYHO	05CHM9_cryptospori
263	30	78.9	332	1	1PXX_HA5IN	P4491 haemophilus	336	30	78.9	3454	2	07Y1G0_ORYSA	07Y1G0_oryza sativ
264	30	78.9	332	1	04OP15_HAE18	04OP15_haemophilus	337	30	78.9	4055	2	06R1L8_CVHNL	06R1L8_cvhnl
265	30	78.9	336	1	SYM_BUCAP	08K941 buchera ap	338	30	78.9	6724	2	06R1L9_CVHNL	06R1L9_cvhnl
266	30	78.9	339	2	030YPO_DESDG	030YPO_desulfovibr	339	30	78.9	6729	2	06Q1S3_CVHNL	06Q1S3_cvhnl
267	30	78.9	341	2	097X18_SULSO	097X18_sulfolobus	340	29	76.3	53	2	08WB30_9NEBP	08WB30_human coron
268	30	78.9	343	2	06ML51_BDEBA	06ML51_bdellovibri	341	29	76.3	53	2	08WBV9_OSTNU	08WBV9_ostriaia fu
269	30	78.9	346	2	041XP6_METBU	041XP6_methanococc	342	29	76.3	67	2	03RK10_RALME	03RK10_ralstonia m
270	30	78.9	348	1	1PXX_AGR15	04uh15 agrobacteri	343	29	76.3	80	2	03P271_9GAMM	03P271_shewanella
271	30	78.9	355	2	04AEF5_SCHLB	04aep5 chlorobium	344	29	76.3	82	2	08VCP7_MOUSE	08VCP7_mus musculu
272	30	78.9	374	2	09KRP2_VIBCH	09KRP2_vibrio chol	345	29	76.3	91	2	08M101_STELE	08M101_centrobacter
273	30	78.9	375	2	04B118_9BURK	04B118_polaromonas	346	29	76.3	99	2	05DBX2_SCHJA	05DBX2_schistosoma
274	30	78.9	386	2	05GRS5_WOLTR	05GRS5_wolbachia s	347	29	76.3	101	2	061G53_DROME	061G53_drosophila
275	30	78.9	396	2	03S0Z6_RALME	03S0Z6_ralstonia m	348	29	76.3	105	2	08A5A8_BACFN	08A5A8_bacteroides
276	30	78.9	400	2	09K1M2_VIBCH	09K1M2_vibrio chol	349	29	76.3	106	2	03TSM4_MOUSE	03TSM4_mus adult mal
277	30	78.9	407	2	03K560_PSEFP	03K560_pseudomonas	350	29	76.3	114	2	02ZSW3_SHEPU	02ZSW3_shewanella
278	30	78.9	407	2	049JF5_ECOLI	049JF5_eshcherichia	351	29	76.3	116	2	03Q1Q1_9GAMM	03Q1Q1_shewanella
279	30	78.9	409	1	MEGFM_PIG	P79385 mus scrofa	352	29	76.3	118	2	02X6K2_9GAMM	02X6K2_shewanella
280	30	78.9	412	2	092Y26_RHIME	092Y26_rhizobium m	353	29	76.3	123	2	02ZRI0_SHEPU	02ZRI0_shewanella
281	30	78.9	413	2	043TL6_SOLUS	043TL6_solibacter	354	29	76.3	119	2	092V24_RHIME	092V24_xanthomonas
282	30	78.9	417	2	057SK8_SALCH	057SK8_salmonella	355	29	76.3	121	2	02PI17_XANOR	02PI17_xanthomonas
283	30	78.9	417	2	05PFK2_SALPA	05PFK2_salmonella	356	29	76.3	123	2	06LY14_METMP	06LY14_methanococc
284	30	78.9	417	2	08ZRG4_SALTY	08ZRG4_salmonella	357	29	76.3	133	2	05GY69_XANOR	05GY69_xanthomonas
285	30	78.9	417	2	08Z914_SALTY	08Z914_salmonella	358	29	76.3	137	2	0939G4_PSEAC	0939G4_pseudomonas
286	30	78.9	424	2	04HC05_9DEIO	04HC05_deinococcus	359	29	76.3	138	2	067N63_SYTHP	067N63_symthobacte
287	30	78.9	427	1	MEGFM_BOVIN	095114 bos taurus	360	29	76.3	139	2	0709A1_AMOGA	0709A1_anopheles g
288	30	78.9	431	2	03T0K7_BOVIN	03T0K7_bos taurus	361	29	76.3	141	2	02UDR7_ASPPR	02UDR7_aespeyillus
289	30	78.9	443	2	03EFU4_ACTSC	03EFU4_actinobacil	362	29	76.3	141	2	05DCR5_SCHJA	05DCR5_schistosoma
290	30	78.9	445	2	044EVO_CHRSL	044EVO_chromohalob	363	29	76.3	142	2	071N46_BACSU	071N46_bacillus su
291	30	78.9	454	2	09RVH3_DEIRA	09RVH3_deinococcus	364	29	76.3	142	2	071N48_BACSU	071N48_bacillus su
292	30	78.9	461	2	043M40_SOLUS	043M40_solibacter	365	29	76.3	144	2	071N50_BACSU	071N50_bacillus su
293	30	78.9	464	2	083X13_SRRPO	083X13_streptomyce	366	29	76.3	146	2	04MFU2_BACC2	04MFU2_bacillus ce
294	30	78.9	467	2	098QJ9_MYCPO	098QJ9_mycoplasma	367	29	76.3	146	2	0631P8_BACC2	0631P8_bacillus ce
295	30	78.9	498	2	082FQ8_STRAW	082FQ8_streptomyce	368	29	76.3	146	2	06HBK8_BACHK	06HBK8_bacillus th
296	30	78.9	501	2	054806_STRPR	054806_streptomyce	369	29	76.3	146	2	081ZK6_BACCI	081ZK6_bacillus ce
297	30	78.9	504	2	06ET11_ORYSA	06ET11_oryza sativ	370	29	76.3	146	2	072Y30_BACCI	072Y30_bacillus ce
298	30	78.9	504	2	03PS41_NITRHO	03PS41_nitrobacter	371	29	76.3	146	2	08RPQ2_BACFN	08RPQ2_bacillus an
299	30	78.9	507	2	09POY3_SRRCO	09POY3_streptomyce	372	29	76.3	157	2	05LFPQ_BACFN	05LFPQ_bacteroides
300	30	78.9	516	2	03HGA3_TRIER	03HGA3_trichodesmi	373	29	76.3	157	2	09FDN5_BACFR	09FDN5_bacteroides
301	30	78.9	526	1	TBC19_HUMAN	08u512 homo sapien	374	29	76.3	158	2	09RYL6_BACTN	09RYL6_bacterioides
302	30	78.9	526	2	09NDX1_HUMAN	09NDX1_homo sapien	375	29	76.3	158	2	09KRW6_VIBCH	09KRW6_vibrio chol
303	30	78.9	527	2	08VDV7_MOUSE	08VDV7_mus musculu	376	29	76.3	168	2	046RX2_RALET	046RX2_ralstonia e
304	30	78.9	527	2	05FNM6_XENLA	05FNM6_xenopus lae	377	29	76.3	171	2	033QY1_9GAMM	033QY1_shewanella
305	30	78.9	528	2	0753S4_ASHGO	0753S4_aeshya goss	378	29	76.3	175	1	DSBB_SHEON	DSBB_shewanella
306	30	78.9	531	2	050UM9_ENTHI	050UM9_entamoeba h	379	29	76.3	175	2	02ZA64_9GAMM	02ZA64_shewanella
307	30	78.9	533	2	04W9P6_ASPTU	04W9P6_aespeyillus	380	29	76.3	175	2	035V85_9GAMM	035V85_shewanella
308	30	78.9	537	2	02UNAN3_ASPPR	02UNAN3_aespeyillus	381	29	76.3	178	2	0367U8_9GAMM	0367U8_shewanella
309	30	78.9	548	1	CUT12_SCHPO	09a9es5 schizosacch	382	29	76.3	186	2	05C112_CRYHO	05C112_cryptospori
310	30	78.9	570	2	04H677_9DEIO	04H677_trypanosoma	383	29	76.3	178	2	03M2Q9_ANAVT	03M2Q9_anabeena va
311	30	78.9	605	2	02UBO9_ASPPR	02UBO9_aespeyillus	384	29	76.3	178	2	065VH4_MANSM	065VH4_mannheimia
312	30	78.9	606	2	Q7VB08_PROMA	Q7VB08_prochloco	385	29	76.3	179	1	GMHB_ANETH	GMHB_aureuriniac
313	30	78.9	616	2	05B8Z5_EMENT	05B8Z5_aespeyillus	386	29	76.3	179	2	Q3Q4F4_9GAMM	Q3Q4F4_shewanella
314	30	78.9	625	2	065D44_BACLD	065D44_bacillus li	387	29	76.3	180	2	Q3N1A4_IDELT	Q3N1A4_synthophoba
315	30	78.9	681	2	09A0E5_9BACN	09a9es5 bacillus sp	388	29	76.3	186	1	PTH_MYCMS	PTH_mycoplasma
316	30	78.9	741	2	038F17_9TRYR	038F17_trypanosoma	389	29	76.3	186	2	02ST16_MYCCA	02ST16_myccoc
317	30	78.9	748	2	07GAL2_ARATH	07GAL2_arabidopsis	390	29	76.3	204	2	08YKT7_ANASP	08YKT7_anabeena sp
318	30	78.9	761	2	09BYJ8_ARATH	09BYJ8_arabidopsis	391	29	76.3	205	2	0418C7_KINRA	0418C7_kineococcus
319	30	78.9	875	2	04QZV4_LEIMA	04QZV4_leishmania	392	29	76.3	205	2	05BV15_SCHJA	05BV15_schistosoma
320	30	78.9	906	2	Q24351_DROME	Q24351_drosophila	393	29	76.3	207	2	Q3XA22_METFL	Q3XA22_methylobaci
321	30	78.9	907	2	Q9VMP4_DROME	Q9VMP4_drosophila	394	29	76.3	218	2	Q96SP2_HUMAN	Q96SP2_homo sapien
322	30	78.9	925	2	Q9ZTAS_ARATH	Q9ZTAS_arabidopsis	395	29	76.3	219	2	Q4E9G8_GRICK	Q4E9G8_wolbachia e
323	30	78.9	974	2	Q3E4U0_CHLAV	Q3E4U0_chloroflexu	396	29	76.3	220	1	Y069_CAEEL	Y069_caenorhabdi

397	29	76.3	221	2	09VZF3_DROME	09vzf3_drosophila	470	29	76.3	335	1	BTUC_PHOTO	07n373_photorhabd
398	29	76.3	221	2	03K4B4_PSEPF	03k4b4_pseudomonas	471	29	76.3	335	1	ITAG2_RAT	053707_tatusu nov
399	29	76.3	223	2	062SR6_BACD	062sr6_bacillus i1	472	29	76.3	335	2	Q53G00_HUMAN	Q53G00_homo sapien
400	29	76.3	223	2	04KJF3_PSEPS	04kjf3_pseudomonas	473	29	76.3	335	2	Q9H0U3_HUMAN	Q9H0U3_homo sapien
401	29	76.3	234	2	07PFU2_ANOGA	07pfu2_anophelis g	474	29	76.3	335	2	Q8NBNG_HUMAN	Q8NBNG_homo sapien
402	29	76.3	235	2	054S81_DICDI	054s81_dicystosteli	475	29	76.3	335	2	Q5RE31_PONPY	Q5re31_pongo pygma
403	29	76.3	237	2	060BR9_PLUXY	060br9_plumella xy	476	29	76.3	335	2	Q9C0Y5_MOUSE	Q9c0y5_m adult mal
404	29	76.3	240	2	09LJ30_HELPS	09lj30_helicobacte	477	29	76.3	335	2	Q6CZT3_MOUSE	Q6cz33_mus musculu
405	29	76.3	241	2	04IHDS_9BACT	04ihds_exiguobacte	478	29	76.3	339	2	Q7B798_PSEAE	Q7b798_p serotype
406	29	76.3	242	2	03W0F8_9ACTO	03w0f8_frankia sp.	479	29	76.3	342	2	Q60ZU2_CAEBR	Q60zu2_caenorhabdi
407	29	76.3	244	2	06SD49_BACD	06sd49_bacillus i1	480	29	76.3	342	2	Q6BDU8_PBRPU	Q6bdu8_periplaneta
408	29	76.3	244	2	06FBX2_ACTIAD	06fbx2_acinetobact	481	29	76.3	342	2	Q2TNRS_9DIP1	Q2tnrs_drosophila
409	29	76.3	244	2	06FC28_ACTIAD	06fc28_acinetobact	482	29	76.3	342	2	Q2TNR6_9DIP1	Q2tnr6_drosophila
410	29	76.3	245	2	02U9D2_ASPOR	02u9d2_aspergillus	483	29	76.3	342	2	Q2TNS9_9DIP1	Q2tns9_drosophila
411	29	76.3	249	2	04WDG5_ASPFU	04wdg5_aspergillus	484	29	76.3	346	2	Q649H8_GARCH	Q649h8_uncultured
412	29	76.3	249	2	02KDM6_RHLET	02kdm6_rhizobium e	485	29	76.3	347	2	Q9SKJ7_ARATH	Q9skj7_arabidops
413	29	76.3	250	2	07RXR2_NEUCR	07rxr2_neurospora	486	29	76.3	348	2	Q59IV6_MESCR	Q59iv6_mesembryant
414	29	76.3	250	2	03WPR7_9RH12	03wpr7_mesorhizobi	487	29	76.3	349	2	Q3E6T0_ARATH	Q3e6t0_arabidops
415	29	76.3	250	2	05FTR3_LACAC	05ftr3_lacobbaci11	488	29	76.3	350	2	Q9SUV2_ARATH	Q9suv2_arabidops
416	29	76.3	251	2	05RCX0_PONPY	05rcx0_pongo pygma	489	29	76.3	351	2	Q9LEH3_ARATH	Q9leh3_arabidops
417	29	76.3	252	1	YQGB_BACSU	P54485_bacillus su	490	29	76.3	352	1	N02M_PARL1	P12771_paracentrot
418	29	76.3	252	2	06SHB1_BACD	06shb1_bacillus i1	491	29	76.3	353	2	Q2S020_9SPH1	Q2s020_salilnibacte
419	29	76.3	256	2	02T4K1_BURRH	02t4k1_burhio fiac	492	29	76.3	354	2	Q6SX78_ORYSA	Q6sx78_oryza sativ
420	29	76.3	256	2	05DZ41_VIBR1	05d41_vibrio fiac	493	29	76.3	356	2	Q6EXK4_LISMO	Q6emx4_listeria mo
421	29	76.3	261	2	0565A7_CAEEL	0565a7_caenorhabd	494	29	76.3	356	2	Q8Y4A1_LISMO	Q8y4a1_listeria mo
422	29	76.3	268	2	Q2K307_RHLET	Q2k307_rhizobium e	495	29	76.3	356	2	Q4S6D4_TETNG	Q4s6d4_tetradon n
423	29	76.3	271	2	Q7X9I6_MALDO	Q7x9i6_malus domes	496	29	76.3	361	2	Q8S2F4_ORYSA	Q8s2f4_oryza sativ
424	29	76.3	273	2	Q7SEFX9_LEPIC	Q7sefx9_leptospira	497	29	76.3	363	2	Q8FHX9_ECOL6	Q8fhx9_escherichia
425	29	76.3	273	2	Q8EXY9_LEPIN	Q8exy9_leptospira	498	29	76.3	365	2	Q397L4_BURS3	Q397l4_burkholderi
426	29	76.3	276	2	Q745N4_MYCPA	Q745n4_mycobacteri	499	29	76.3	367	2	Q4BLV0_BURVI	Q4blv0_burkholderi
427	29	76.3	276	2	Q9CDH2_MYCLE	Q9cdh2_mycobacteri	500	29	76.3	368	2	Q2U1J3_ASPOR	Q2u1j3_aspergillus
428	29	76.3	278	2	Q2IFC7_9DELT	Q2ifc7_aeromonxob							
429	29	76.3	279	2	02S483_HELPS	02s483_helicobacte							
430	29	76.3	279	2	09ZL41_HELPS	09z141_helicobacte							
431	29	76.3	283	2	05GRU6_WOLRR	05gru6_wolbachia s							
432	29	76.3	284	2	035DE7_9BRAD	035de7_brayirrhizoc							
433	29	76.3	285	2	Q3DTD7_STRAG	Q3dtd7_streptococc							
434	29	76.3	290	2	059078_ALTRAT	059078_alteromonas							
435	29	76.3	290	2	06RW45_9GAMM	06rw45_pseudalter							
436	29	76.3	299	2	05LA07_BACFN	05la07_bacteroides							
437	29	76.3	299	2	0640B6_BACFR	0640b6_bacteroides							
438	29	76.3	300	2	0257G6_9GAMM	02s7g6_hahella che							
439	29	76.3	300	2	Q3F7E0_9BURK	Q3f7e0_burkholderi							
440	29	76.3	300	2	Q45DK1_9BURK	Q45dk1_burkholderi							
441	29	76.3	300	2	Q4LWP5_9BURK	Q4lwp5_burkholderi							
442	29	76.3	301	1	REIS_TODPA	P23820_coderodes p							
443	29	76.3	301	2	05RBG3_PONPY	05rbg3_pongo pygma							
444	29	76.3	301	2	08XRN9_TALSO	08xrn9_faltonoma s							
445	29	76.3	302	1	CITG_KLEPN	P45414_klebsiella s							
446	29	76.3	302	2	Q73IG4_WOLPM	Q73194_wolbachia p							
447	29	76.3	303	2	060R86_CAEBR	060r86_caenorhabdi							
448	29	76.3	303	2	03OJAI_9GAMM	03ojai_shevanella							
449	29	76.3	303	2	051364_PSEAE	051364_pseudomonas							
450	29	76.3	303	2	05DOH1_ECOLI	05doh1_escherichia							
451	29	76.3	304	2	05VCE2_9SAUR	05vc2_rieppelion							
452	29	76.3	305	2	05VCE3_9SAUR	05vc3_rieppelion							
453	29	76.3	305	2	05VCP4_9SAUR	05vc4_rieppelion							
454	29	76.3	305	2	05VCP5_9SAUR	05vc5_rieppelion							
455	29	76.3	305	2	05VCP6_9SAUR	05vc6_rieppelion							
456	29	76.3	306	2	05LMA1_SILPO	05lma1_silicobacte							
457	29	76.3	310	2	Q332K2_RHOPA	Q332k2_rhodospande							
458	29	76.3	312	2	06F6C2_9ACAR	06f6c2_ixodes uria							
459	29	76.3	313	2	027708_NEPNO	027708_nephtros no							
460	29	76.3	315	2	0871X0_VIBPA	0871x0_vibrio para							
461	29	76.3	319	2	02UDV1_ASPOR	02udv1_aspergillus							
462	29	76.3	319	2	04HPC0_CAMPD	04hpc0_campylobact							
463	29	76.3	320	2	Q36AF9_9GAMM	Q36af9_hemewanella							
464	29	76.3	321	1	CYSP3_HOMAM	P25764_homatus ame							
465	29	76.3	323	2	04SFMG_TETNG	04sfmg_tetradon n							
466	29	76.3	325	2	048375_BPPI	048375_bacteriopa							
467	29	76.3	326	2	04E7N8_BRICK	04e7n8_wolbachia e							
468	29	76.3	328	2	09T9N2_9BILA	09t9n2_tetradualli							
469	29	76.3	331	2	05EXH9_MACRS	05exh9_macrobrachi							

RESULT 1

ID Q8AGE6 CORGL PRELIMINARY; PRT; 71 AA.

AC Q8AGE6;

DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.

DT 01-JUN-2003, sequence version 1.

DT 07-FEB-2006, entry version 11.

DE Hypothetical protein.

OS Corynebacterium glutamicum (Brevibacterium flavum).

OG Plasmid pC32.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacteriaceae; Corynebacterium.

OX NCBI_TaxId=1718;

OX NCU007IDE SEQUENCE.

RP STRAIN=ATCC 31832;

RC MEDLINE=22830013; PubMed=12948627; DOI=10.1016/S0168-1656(03)00157-3;

RA Tauch A., Puhler A., Kainowski J., Thierbach G.;

RT "Plasmids in Corynebacterium glutamicum and their molecular classification by comparative genomics.";

RL J. Biotechnol. 104:27-40(2003).

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>

CC Distributed under the Creative Commons Attribution-NonDerivs License

CC EMBL, AY172685; AA018202.1; -; Genomic DNA.

KW Hypothetical protein; Plasmid.

SQ SEQUENCE 71 AA; 7742 MW; DF8FD8708D007D11 CRC64;

Query Match

Best Local Similarity 92.1%; Score 35; DB 2; Length 71;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 WMLSAFS 7

58 MWVSASF 64

RESULT 2

Q8N0Z0_HUMAN PRELIMINARY; PRT; 347 AA.
 ID Q8N0Z0_HUMAN
 AC Q8N0Z0;
 DT 01-OCT-2002, integrated into UniProtKB/TrEMBL.
 DT 01-OCT-2002, sequence version 1.
 DT 07-FEB-2006, entry version 15.
 DE Seven transmembrane helix receptor.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 NCBI_TaxID=9606;
 [1]
 NP NUCLEOTIDE SEQUENCE.
 RA Suwa M., Sato T., Okouchi I., Arita M., Futami K., Matsumoto S.,
 RA Tezuka S., Aburatani H., Asai K., Akiyama Y.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs License

 DR EMBL: AB065450; BAC05713.1; -; Genomic DNA.
 DR EMBL: AB065566; BAC05802.1; -; Genomic DNA.
 DR EMBL: AB065486; BAC05739.1; -; Genomic DNA.
 DR Ensembl: ENSG00000198822; Homo sapiens.
 DR GO: GO:0016021; C: integral to membrane; IEA.
 DR GO: GO:0008067; F: metabotropic glutamate, GABA-B-like recepto. .; IEA.
 DR GO: GO:0004872; F: receptor activity; IEA.
 DR InterPro: IPR000337; GPCR_Mgt.
 DR InterPro: IPR011500; NCD3G_GPCR.
 DR Pfam: PF070562; NCD3G_1.
 DR PRINTS: PR00248; GPCRMR.
 DR PROSITE: PS00980; G_PROTEIN_RECEP_F3_2; 1.
 DR PROSITE: PS00981; G_PROTEIN_RECEP_F3_3; 1.
 DR PROSITE: PS50259; G_PROTEIN_RECEP_F3_4; 1.
 DR Receptor; Transmembrane.
 SQ SEQUENCE 347 AA; 39300 MW; B114CF38477FE624 CRC64;

Query Match 92.1%; Score 35; DB 2; Length 347;
 Best Local Similarity 85.7%; Pred. No. 75;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WMLSAFS 7
 |||||:
 Db 339 WMLSSFS 345

RESULT 3

Q5L241_GEOKA PRELIMINARY; PRT; 424 AA.
 ID Q5L241_GEOKA
 AC Q5L241;
 DT 01-FEB-2005, integrated into UniProtKB/TrEMBL.
 DT 01-FEB-2005, sequence version 1.
 DT 07-FEB-2006, entry version 6.
 DE Maltose/maltodextrin transport system (Substrate-binding protein).
 GN Oederlucuskaues-GK0704;
 OS Geobacillus kaustophilus.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Geobacillus.
 NCBI_TaxID=1462;
 [1]
 NP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=HTF426;
 RX PubMed=15576355; DOI=10.1093/nar/gkh970;
 RA Takami H., Takaki Y., Chee G.-J., Nishi S., Shimanura S., Suzuki H.,
 RA Matsui S., Uchiyama I.;
 RT "Thermoadaptation trait revealed by the genome sequence of
 RT thermophilic Geobacillus kaustophilus.";
 RL Nucleic Acids Res. 32:6292-6303(2004).

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs License

DR EMBL: BA00043; BAD74989.1; -; Genomic DNA.
 DR GO: GO:0005363; F: maltose transporter activity; IEA.
 DR GO: GO:0005215; F: transporter activity; IEA.
 DR GO: GO:0015768; P: maltose transport; IEA.
 DR GO: GO:0006810; P: transport; IEA.
 DR InterPro: IPR006060; Maltose_bd.
 DR InterPro: IPR006059; SBP_bac_1.
 DR InterPro: IPR006061; SBP_dom1.
 DR Pfam: PF01547; SBP_bac_1; 1.
 DR PRINTS: PR00181; MALTOSBP.
 DR PROSITE: PS01037; SBP_BACTERIAL_1; UNKNOWN_1.
 KM Complete proteome.
 SQ SEQUENCE 424 AA; 46652 MW; E16D1884B5A57B6E CRC64;

Query Match 92.1%; Score 35; DB 2; Length 424;
 Best Local Similarity 85.7%; Pred. No. 92;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WMLSAFS 7
 |||||:
 Db 310 WMLSAFS 316

RESULT 4

Q9BEF3_THRSW PRELIMINARY; PRT; 459 AA.
 ID Q9BEF3_THRSW
 AC Q9BEF3;
 DT 01-JUN-2001, integrated into UniProtKB/TrEMBL.
 DT 01-JUN-2001, sequence version 1.
 DT 07-FEB-2006, entry version 19.
 DE NADH-ubiquinone oxidoreductase chain 4 (EC 1.6.5.3) (NADH
 DE dehydrogenase subunit 4).
 GN Name=NADH4;
 OS Thryonoms swinderianus (Greater cane rat).
 OC Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 OC Hystriocognathi; Thryonomysidae; Thryonomys.
 NCBI_TaxID=10169;
 [1]
 NP NUCLEOTIDE SEQUENCE.
 RP MEDLINE=2109735; PubMed=11161749; DOI=10.1006/mpev.2000.0870;
 RA Mouchaty S.K., Catzeffis F., Janke A., Arnason U.;
 RT "Molecular evidence of an African Phiomorpha-South American
 RT Cavimorpha clade and support for Hystriocognathi based on the complete
 RT mitochondrial genome of the cane rat (Thryonomys swinderianus).";
 RL Mol. Phylogenet. Evol. 18:127-135(2001).
 CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
 CC -1- SIMILARITY: Belongs to the complex I subunit 4 family.
 CC
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs License

 DR EMBL: AJ301644; CAC27808.1; -; Genomic DNA.
 DR GO: GO:0005739; C: mitochondrion; IEA.
 DR GO: GO:0008137; F: NADH dehydrogenase (ubiquinone) activity; IEA.
 DR GO: GO:0016491; F: oxidoreductase activity; IEA.
 DR GO: GO:0042773; P: ATP synthesis coupled electron transport; IEA.
 DR GO: GO:0006120; P: mitochondrial electron transport; IEA.
 DR InterPro: IPR003918; NADHbd_oxred4.
 DR InterPro: IPR010227; NDH_I_M.
 DR InterPro: IPR001750; Oxidored_q1.
 DR InterPro: IPR000260; Oxidored_q5_N.
 DR PANTHER: PTHR11460; NADHbd_oxred4; 1.
 DR Pfam: PF00361; Oxidored_q1_1.
 DR Pfam: PF01059; Oxidored_q5_N_1.
 DR PRINTS: PR01437; NDOXDDBTSE4.
 DR TIGRFAMs: TIGR01972; NDH_I_M_1.
 KM Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
 SQ SEQUENCE 459 AA; 51966 MW; E6DD0150CC0C9F56 CRC64;

Query Match 92.1%; Score 35; DB 2; Length 459;
 Best Local Similarity 85.7%; Pred. No. 99;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WMLSAFS 7
 Db 359 WMLSSFS 365

RESULT 5
 Q9VLT3 DROME PRELIMINARY; PRT; 1760 AA.
 ID Q9VLT3 DROME
 AC Q9VLT3;
 DT 01-MAY-2000, integrated into UniProtKB/TrEMBL.
 DT 01-MAY-2000, sequence version 1.
 DT 07-FEB-2006, entry version 32.
 DE CG7586-PA (L023292P).
 GN Name=Mcr; ORFNames=CG7586; Dmel CG7586;
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 CC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 CC Ephydroidae; Drosophilidae; Drosophila.
 OC NCBI_taxid=7227;
 OX [1]

NP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champs M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bertan B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan P.V., Bouck J., Brokstein P., Brotlier P.,
 RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriere S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jatali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kratz C., Kravitz S., Kuip D., Lai Z.,
 RA Laake P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milhina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puti V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kimms I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskaas R., Tecor C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 RN [2]

RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22426065; PubMed=12537568;
 RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
 RA Patel S., Adams M., Champs M., Dugan S.P., Frise E., Hodgson A.,
 RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,
 RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
 RA Svirskaas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,

RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
 RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
 RT melanogaster euchromatic genome sequence.";
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).

RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22426070; PubMed=12537573;
 RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirskaas R.,
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
 RA Ashburner M., Celniker S.E.;
 RT "The transposable elements of the Drosophila melanogaster euchromatin:
 RT a genome perspective.";
 RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).

RN [4]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22426069; PubMed=12537572;
 RA Mira S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bertan B.P.,
 RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.O.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a
 RT systematic review.";
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).

RN [5]
 RP NUCLEOTIDE SEQUENCE.
 RG Berkeley Drosophila Genome Project;
 RA Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
 RA Hoskins R., Stapleton M., Pacleb J., Park S., Svirskaas R., Smith E.,
 RA Yu C., Rubin G.;
 RT "Drosophila melanogaster release 4 sequence.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RN [6]
 RP NUCLEOTIDE SEQUENCE.
 RG FlyBase;
 RL Submitted (JAN-2006) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP NUCLEOTIDE SEQUENCE.
 RA STRAIN=Berkeley.
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champs M., Chaver C., Dorsett V., Farfan D., Frise E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonCommercial License
 CC -----
 DR EMBL: AE003619; AAF52601.1; -; Genomic DNA.
 DR EMBL: AY058526; AAL13755.1; -; mRNA.
 DR HSSP: P01130; 1A0J.

DR InFact; Q9VLT3; -;
 DR FLYBase; FBgn0020240; MCr.
 DR GO: GO:0005515; P:protein binding; IPI.
 DR InterPro: IPR011626; A2M comp.
 DR InterPro: IPR002890; A2M_N.
 DR InterPro: IPR011625; A2M_N_2.
 DR InterPro: IPR011627; A2M_icpt.
 DR InterPro: IPR002172; LDL_icpt_A.
 DR InterPro: IPR001599; Macroglublna2.
 DR Pfam: PF00207; A2M; 1.
 DR Pfam: PF07678; A2M_comp; 1.
 DR Pfam: PF01835; A2M_N; 1.
 DR Pfam: PF07703; A2M_N_2; 1.
 DR Pfam: PF07677; A2M_icpt; 1.
 DR Pfam: PF00057; ldl_icpt_a; 1.
 DR SMART: SM00192; LDLa; 1.
 DR PROSITE: PS50068; LDLRA_2; 1.

DR PROSITE: PS50068; LDLRA_2; 1.
 DR SEQUENCE 1760 AA; 202783 MW; B8E0602350EB87F86 CRC64;

```

Query Match      92.1%; Score 35; DB 2; Length 1760;
Best Local Similarity 85.7%; Pred. No. 4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 WMLSAFS 7
      ||:|||||
Db      966 WMLSAFS 972

RESULT 6
Q34A91_RHOA      PRELIMINARY; PRT; 2943 AA.
ID Q34A91_RHOA
AC Q34A91
DT 06-DEC-2005, integrated into UniProtKB/TrEMBL.
DT 06-DEC-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Amino acid acetylation.
GN ORFNames=RPCDRAFT_2142;
OS Rhodopseudomonas palustris B12B18.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiales; Rhodopseudomonas.
OX NCBI_TaxID=316056;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=B12B18;
RG US DOE Joint Genome Institute (JGI-DOE);
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter J.C., Glavina T.,
RA Hammon N., Istrail S., Pitluck S., Richardson P.;
RT "Sequencing of the draft genome and assembly of Rhodopseudomonas
RT palustris B12B18."
RL Submitted (OCT-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=B12B18;
RG US DOE Joint Genome Institute (JGI-DOE);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Rhodopseudomonas palustris
RT B12B18."
RL Submitted (NOV-2005) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
-----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC
CC EMBL; AALR0100008; EAP1865.1; -; Genomic_DNA.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0030170; F:pyridoxal phosphate binding; IEA.
DR GO; GO:0008483; F:transaminase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006633; P:fatty acid biosynthesis; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
SQ SEQUENCE 2943 AA; 308266 MW; 358AE8F58262601C CRC64;

Query Match      92.1%; Score 35; DB 2; Length 2943;
Best Local Similarity 85.7%; Pred. No. 6.7e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 WMLSAFS 7
      ||:|||||
Db      1502 WMLSAFS 1508

RESULT 7
Q4MRP7_BACCE      PRELIMINARY; PRT; 147 AA.
ID Q4MRP7_BACCE
AC Q4MRP7
DT 02-AUG-2005, integrated into UniProtKB/TrEMBL.
DT 02-AUG-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Non-specific DNA-binding protein Dps / iron-binding ferritin-like
DE antioxidant protein / ferroxidase.

```

```

GN Name=dps; ORFNames=BCE_G9241_2031;
OS Bacillus cereus G9241.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC Bacillus cereus group.
OX NCBI_TaxID=269801;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=G9241;
RG PubMed=15155910; DOI=10.1073/pnas.0402414101;
RA Hofmaster A.R., Ravel J., Rasko D.A., Chapman G.D., Chute M.D.,
RA Marston C.K., De B.K., Sacchi C.T., Fitzgerald C., Mayer L.W.,
RA Malden M.C.J., Priest F.G., Barker M., Jiang L., Cer R.Z.,
RA Ristone J., Peterson S.N., Weyant R.S., Galloway D.R., Read T.D.,
RA Popovic T., Fraser C.M.;
RT "Identification of anthrax toxin genes in a Bacillus cereus associated
RT with an illness resembling inhalation anthrax."
RL Proc. Natl. Acad. Sci. U.S.A. 101:8449-8454 (2004).
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
-----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC
CC EMBL; AAEK0100011; EML14754.1; -; Genomic_DNA.
DR SMR; Q4MRP7; 2-147
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008199; F:ferric iron binding; IEA.
DR GO; GO:0006879; P:iron ion homeostasis; IEA.
DR GO; GO:0006950; P:response to stress; IEA.
DR InterPro; IPR02177; DPS_DNA_bd.
DR Pfam; PF00210; Ferritin; 1.
DR PRINTS; PR01346; HELMADAPROT.
DR PRODOM; PD149803; DPS; 1.
DR PROSITE; PS00818; DPS_1; 1.
KW DNA-binding.
SQ SEQUENCE 147 AA; 16650 MW; D8A5654D7E0D77CD CRC64;

Query Match      89.5%; Score 34; DB 2; Length 147;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WMLSAF 6
      |||||
Db      140 WMLSAF 145

RESULT 8
Q63CE7_BACCC      PRELIMINARY; PRT; 147 AA.
ID Q63CE7_BACCC
AC Q63CE7
DT 25-OCT-2004, integrated into UniProtKB/TrEMBL.
DT 25-OCT-2004, sequence version 1.
DT 07-FEB-2006, entry version 14.
DE Non-specific DNA-binding protein. Dps-like (EC 1.16.3.1).
GN Name=dps; OrderedlocusNames=BCEJ311825;
OS Bacillus cereus (strain ZK / E33L).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC Bacillus cereus group.
OX NCBI_TaxID=288681;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RA Bretin T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K.,
RA Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,
RA Richardson P., Rubin E., Tice H.;
RT "Complete genome sequence of Bacillus cereus ZK."
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC
CC EMBL; CP000001; AAU18426.1; -; Genomic_DNA.
DR SMR; Q63CE7; 2-147.

```

```

DR GO: GO:0003677; F:DNA binding; IEA.
DR GO: GO:0008199; F:ferric iron binding; IEA.
DR GO: GO:0004322; F:ferroxidase activity; IEA.
DR GO: GO:0016491; F:oxidoreductase activity; IEA.
DR GO: GO:0006879; P:iron ion homeostasis; IEA.
DR GO: GO:0006950; P:response to stress; IEA.
DR InterPro: IPR002177; DPS_DNA_bd.
DR InterPro: IPR008331; Ferritin_Dps.
DR InterPro: IPR012347; Ferritin_rel.
DR Pfam: PF00210; Ferritin_1.
DR PRINTS: PR01346; HELMAPAPROT.
DR ProDom: PD149803; DPS_1.
DR PROSITE: PS00818; DPS_1.
DR Complete proteome: DNA-binding; Oxidoreductase.
SQ SEQUENCE 147 AA; 16707 MW; 2745255C84FCCCCD CRC64;

Query Match 89.5%; Score 34; DB 2; Length 147;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 WMLSAF 6
Db 140 WMLSAF 145

RESULT 9
06HJV3_BACCH PRELIMINARY; PRT; 147 AA.
ID 06HJV3_BACCH
AC 06HJV3
DT 19-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, entry version 14.
DE Non-specific DNA-binding protein Dps (EC 1.16.3.1).
GN Name=dps; OrderedLocNames=BT9727.1841;
OS Bacillus thuringiensis subsp. konkukian.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC Bacillus cereus group.
OX NCBI_TaxID=180856;
[1]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP STRAIN=97-27;
RA Bretin T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K.,
RA Hitchcock P., Jackson P., Kelm P., Longmire J., Lucas S., Okinaka R.,
RA Richardson P., Rubin E., Rice H.;
RT "Complete genome sequence of Bacillus thuringiensis 97-27.";
RL Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL: AE017355; AAT63428.1; -; Genomic_DNA.
DR SMR: 06HJV3; 2-147.
DR GO: GO:0003677; F:DNA binding; IEA.
DR GO: GO:0008199; F:ferric iron binding; IEA.
DR GO: GO:0004322; F:ferroxidase activity; IEA.
DR GO: GO:0016491; F:oxidoreductase activity; IEA.
DR GO: GO:0006879; P:iron ion homeostasis; IEA.
DR GO: GO:0006950; P:response to stress; IEA.
DR InterPro: IPR002177; DPS_DNA_bd.
DR InterPro: IPR008331; Ferritin_Dps.
DR InterPro: IPR012347; Ferritin_rel.
DR Pfam: PF00210; Ferritin_1.
DR PRINTS: PR01346; HELMAPAPROT.
DR ProDom: PD149803; DPS_1.
DR PROSITE: PS00818; DPS_1.
DR Complete proteome: DNA-binding; Oxidoreductase.
SQ SEQUENCE 147 AA; 16707 MW; 2745255C84FCCCCD CRC64;

Query Match 89.5%; Score 34; DB 2; Length 147;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 WMLSAF 6
Db 140 WMLSAF 145

```

```

Db 140 WMLSAF 145

RESULT 10
0813P2_BACCR PRELIMINARY; PRT; 147 AA.
ID 0813P2_BACCR
AC 0813P2
DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2003, sequence version 1.
DT 07-FEB-2006, entry version 18.
DE Non-specific DNA-binding protein Dps / Iron-binding ferritin-like
DE antioxidant protein / Ferroxidase (EC 1.16.3.1).
GN OrderedLocNames=BC2011; ORFNames=BC_2011.
OS Bacillus cereus (strain ATCC 14579 / DSM 31).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC Bacillus cereus group.
OX NCBI_TaxID=226900;
[1]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP MEDLINE=22608415; PubMed=12721630; DOI=10.1038/nature01582;
RX Ivanova N., Sorokin A., Anderson I., Galleron N., Candelon B.,
RA Kaparatel V., Bhattacharya A., Reznik G., Mikhailova N., Lapidus A.,
RA Chu L., Mazur M., Goltzman E., Larsen N., D'Souza M., Malinas T.,
RA Grechkin Y., Pusch G., Haselkorn R., Fomstein M., Ehtlich S.D.,
RA Overbeek R., Kyriides N.C.;
RT "Genome sequence of Bacillus cereus and comparative analysis with
RT Bacillus anthracis.";
RL Nature 423:87-91(2003).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL: AE016877; AAP08980.1; -; Genomic_DNA.
DR HSSP: Q8RP01; 141G.
DR SMR: 0813P2; 2-147.
DR Biocyc: BCR226900.BC2011-MONOMER; -.
DR GO: GO:0003677; F:DNA binding; IEA.
DR GO: GO:0008199; F:ferric iron binding; IEA.
DR GO: GO:0004322; F:ferroxidase activity; IEA.
DR GO: GO:0016491; F:oxidoreductase activity; IEA.
DR GO: GO:0006879; P:iron ion homeostasis; IEA.
DR GO: GO:0006950; P:response to stress; IEA.
DR InterPro: IPR002177; DPS_DNA_bd.
DR InterPro: IPR008331; Ferritin_Dps.
DR InterPro: IPR012347; Ferritin_rel.
DR Pfam: PF00210; Ferritin_1.
DR PRINTS: PR01346; HELMAPAPROT.
DR ProDom: PD149803; DPS_1.
DR PROSITE: PS00818; DPS_1.
DR Complete proteome: DNA-binding; Oxidoreductase.
SQ SEQUENCE 147 AA; 16680 MW; 36A4D54D650CCCCD CRC64;

Query Match 89.5%; Score 34; DB 2; Length 147;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 WMLSAF 6
Db 140 WMLSAF 145

RESULT 11
0739P8_BACCI PRELIMINARY; PRT; 147 AA.
ID 0739P8_BACCI
AC 0739P8
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-MAR-2006, entry version 14.
DE General stress protein.
GN Name=dps; ORFNames=BCE_2092;
OS Bacillus cereus (strain ATCC 10987).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;

```

```

OC Bacillus cereus group.
OX NCBI_TaxID=222523;

[1]
NP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP PubMed=1960714; DOI=10.1093/nar/gkh258;
RA Rasko D.A., Ravel J., Oekstad O.A., Helgason E., Cer R.Z., Jiang L.,
RA Shores K.A., Fouts D.E., Tourasse N.J., Angiloul S.V., Kolonay J.F.,
RA Nelson W.C., Koliste A.-B., Fraser C.M., Read T.D.;
RT "The genome sequence of Bacillus cereus ATCC 10987 reveals metabolic
RT adaptations and a large plasmid related to Bacillus anthracis pXO1."
RL Nucleic Acids Res. 32:977-988(2004).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL; AE017194; AAS41013.1; -; Genomic_DNA.
DR SMR; Q739P8; 2-147.
DR GenomeReviews; AE017194.GR; dps.
DR GO; GO:0008199; F:ferric iron binding; IEA.
DR GO; GO:0006950; P:response to stress; IEA.
DR GO; GO:0006950; P:response to stress; IEA.
DR InterPro; IPR002177; DPS_DNA_bd.
DR InterPro; IPR008331; Ferritin_Dps.
DR Pfam; PF00210; Ferritin; 1.
DR PRINTS; PR01346; HEINAPAPROT.
DR PRODOM; PD149803; DPS; 1.
DR PROSITE; PS00818; DPS_1; 1.
KM Complete proteome.
SQ SEQUENCE 147 AA; 16707 MW; 2745255C84FCCCD CRC64;

Query Match 89.5%; Score 34; DB 2; Length 147;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WMLSAF 6
Db 140 WMLSAF 145

RESULT 12
Q8RP01 BACAN PRELIMINARY; PRT; 147 AA.
AC Q8RP01; Q6HZV3; Q6KTT5; Q8IKM9;
DT 01-JUN-2002, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2002, sequence version 1.
DT 07-MAR-2006, entry version 25.
DE Mini-ferritin Dlp2 (General stress protein).
GN Name=dlp2; Synonyms=dps; OrderedLocustNames=BA2013, BAS1871, GBA2013;
GN ORFNames=BA_2013;
OS Bacillus anthracis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC Bacillus cereus group.
OX NCBI_TaxID=1392;

[1]
NP NUCLEOTIDE SEQUENCE.
RP MEDLINE=21964043; PubMed=11836250; DOI=10.1074/jbc.M112378200;
RA Papiunto E., Dundon W.G., Piculis N., Battistutta R., Montecucco C.,
RA Zanotti G.;
RT "Structure of two iron-binding proteins from Bacillus anthracis."
RL J. Biol. Chem. 277:15093-15098(2002).

[2]
NP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP STRAIN=Ames / isolate Porton;
RX MEDLINE=22608414; PubMed=12721629; DOI=10.1038/nature01586;
RT 08-NOV-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Non-specific DNA-binding protein Dps / Iron-binding ferritin-like
DE antioxidant protein / Peroxidase (EC 1.16.3.1).
GN ORFNames=BTH_06140;
OS Bacillus thuringiensis serovar israelensis ATCC 35646.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC Bacillus cereus group.
OX NCBI_TaxID=339854;
[1]

```

```

RA Hazen A., Cline R.T., Redmond C., Thwaitte J.E., White O.,
RA Salzberg S.L., Thomson B., Friedlander A.M., Koehler T.M.,
RA Hanna P.C., Koliste A.-B., Fraser C.M.;
RT "The genome sequence of Bacillus anthracis Ames and comparison to
RT closely related bacteria."
RL Nature 423:81-86(2003).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL; AF374269; AAM18636.1; -; Genomic_DNA.
DR EMBL; AE016879; AAP25903.1; -; Genomic_DNA.
DR EMBL; AE017225; AAT54186.1; -; Genomic_DNA.
DR EMBL; AE017334; AAT31135.1; -; Genomic_DNA.
DR PDB; 1JTG; X-ray; A/B/C/D=2-147.
DR GenomeReviews; AE016879.GR; BA2013.
DR GenomeReviews; AE017225.GR; BAS1871.
DR GenomeReviews; AE017334.GR; GBA2013.
DR TIGR; BA2013; -.
DR TIGR; GBA2013; -.
DR GO; GO:0008199; F:ferric iron binding; IEA.
DR GO; GO:0006979; P:iron ion homeostasis; IEA.
DR GO; GO:0006950; P:response to stress; IEA.
DR InterPro; IPR002177; DPS_DNA_bd.
DR InterPro; IPR008331; Ferritin_Dps.
DR InterPro; IPR01347; Ferritin_rel.
DR Pfam; PF00210; Ferritin; 1.
DR PRINTS; PR01346; HEINAPAPROT.
DR PRODOM; PD149803; DPS; 1.
DR PROSITE; PS00818; DPS_1; 1.
KM Complete proteome.
SQ SEQUENCE 147 AA; 16649 MW; 2741651884FCCCD CRC64;

Query Match 89.5%; Score 34; DB 2; Length 147;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WMLSAF 6
Db 140 WMLSAF 145

RESULT 13
Q3EXQ4 BACTI PRELIMINARY; PRT; 154 AA.
ID Q3EXQ4; BACTI
AC Q3EXQ4;
DT 08-NOV-2005, integrated into UniProtKB/TrEMBL.
DT 08-NOV-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Non-specific DNA-binding protein Dps / Iron-binding ferritin-like
DE antioxidant protein / Peroxidase (EC 1.16.3.1).
GN ORFNames=BTH_06140;
OS Bacillus thuringiensis serovar israelensis ATCC 35646.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC Bacillus cereus group.
OX NCBI_TaxID=339854;
[1]

```

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=ATCC 35646;
 RA Anderson I., Sorokin A., Kapral V., Reznik G., Bhattacharya A.,
 RA Mikhailova N., Burd H., Joukov V., Kaznadzey D., Walunas T.,
 RA D'Souza M., Larsen N., Pusch G., Lioios K., Greckin Y., Lapidas A.,
 RA Goldsman E., Chu L., Feinstein M., Ehrlich D., Overbeek R.,
 RA Kyrpides N., Ivanova N.;
 RT "Comparative genome analysis of *Bacillus cereus* group genomes with
 RT *Bacillus subtilis*.";
 RL Submitted (SEP-2005) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs License
 CC -----
 DR EMBL, AAJ01000037; EA056104.1; -; Genomic_DNA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0008199; F:ferri iron binding; IEA.
 DR GO; GO:0004322; F:ferroxidase activity; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006879; P:iron ion homeostasis; IEA.
 DR GO; GO:0006950; P:response to stress; IEA.
 DR InterPro; IPR002177; Dfs_DNA_Bd.
 DR InterPro; IPR008331; Ferritin_Dfs.
 DR InterPro; IPR012347; Ferritin_rel.
 DR Pfam; PF00210; Ferritin_1.
 DR PRINTS; PR01346; HELNAPAPROT.
 DR PRODOM; PD149803; Dfs; 1.
 DR PROSITE; PS00818; Dfs_1; 1.
 DR DNA-binding; Oxidoreductase.
 SQ SEQUENCE 154 AA; 17397 MW; 1C179A213C87D3C0 CRC64;

Query Match 89.5%; Score 34; DB 2; Length 154;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WMLSAF 6
 |||||
 Db 147 WMLSAF 152

RESULT 14
 Q5E524 VIBF1 PRELIMINARY; PRT; 211 AA.
 AC Q5E524;
 DT 15-MAR-2005, integrated into UniProtKB/TrEMBL.
 DT 07-FEB-2006, entry version 1.
 DE Lysine exporter protein.
 GN OrderedLocustNames=VF1057;
 OS *Vibrio fischeri* (strain ATCC 700601 / ES114).
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 CC Vibrionaceae; Vibrrio.
 CC NCBI_TaxID=312309;
 CC [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RX PubMed=15703294; DOI=10.1073/pnas.0409900102;
 RA Ruby E.G., Urbanowski M., Campbell J., Dunn A., Faini M., Gonsalus R.,
 RA Loshron P., Lupp C., McCam J., Millikan D., Schaefer A., Stabb E.,
 RA Stevens A., Vaisick K., Whistler C., Greenberg E.P.;
 RT "Complete genome sequence of *Vibrio fischeri*: a symbiotic bacterium
 RT with pathogenic congeners.";
 RL Proc. Natl. Acad. Sci. U.S.A. 102:3004-3009(2005).
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs License
 CC -----
 DR EMBL, CP000020; AAM8552.1; -; Genomic_DNA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005293; F:lysine permease activity; IEA.
 DR GO; GO:0006865; P:amino acid transport; IEA.

DR InterPro; IPR001123; Lyse.
 DR Pfam; PF01810; Lyse; 1.
 KW Complete proteome.
 SQ SEQUENCE 211 AA; 22582 MW; F69806C1F7B1D5D9 CRC64;

Query Match 89.5%; Score 34; DB 2; Length 211;
 Best Local Similarity 100.0%; Pred. No. 73;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WMLSAF 6
 |||||
 Db 203 WMLSAF 208

RESULT 15
 Q744H1 MYCPA PRELIMINARY; PRT; 361 AA.
 ID Q744H1_MYCPA
 AC Q744H1;
 DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
 DT 05-JUL-2004, sequence version 1.
 DT 07-FEB-2006, entry version 10.
 DE Hypothetical protein.
 GN OrderedLocustNames=MAP0370; ORFNames=MAP_0370;
 OS *Mycobacterium paratuberculosis*.
 CC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 CC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
 CC *Mycobacterium avium* complex (MAC).
 CC NCBI_TaxID=1770;
 CC [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=K10;
 RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs License
 CC -----
 DR EMBL, AE016958; AAS02687.1; -; Genomic_DNA.
 DR InterPro; IPR005614; NtFD.
 DR Pfam; PF03916; NtFD; 1.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 361 AA; 37516 MW; 9AACB052006E772C CRC64;

Query Match 89.5%; Score 34; DB 2; Length 361;
 Best Local Similarity 85.7%; Pred. No. 1,36+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
 |||||
 Db 138 WMLSAFS 144

RESULT 16
 Q2W9C2 MAGSA PRELIMINARY; PRT; 458 AA.
 ID Q2W9C2_MAGSA
 AC Q2W9C2;
 DT 10-JAN-2006, integrated into UniProtKB/TrEMBL.
 DT 10-JAN-2006, sequence version 1.
 DT 07-MAR-2006, entry version 4.
 DE Permease of the major facilitator superfamily.
 GN ORFNames=amb0749;
 OS *Magnetospirillum magneticum* AMB-1.
 CC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
 CC Rhodospirillaceae; Magnetospirillum.
 CC NCBI_TaxID=342108;
 CC [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=AMB-1;
 RX PubMed=16303747;
 RA Matsunaga T., Okamura Y., Fukuda Y., Mahyudi A.T., Murase Y.,
 RA Takeyama H.;
 RT "Complete Genome Sequence of the Facultative Anaerobic Magnetotactic
 RT Bacterium *Magnetospirillum* sp. strain AMB-1.";


```

RL  DNA Res. 12:157-166(2005).
CC  -|- SUBCELLULAR LOCATION: Inner membrane; multi-pass membrane protein
CC  (by similarity).
CC  -----
CC  Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC  Distributed under the Creative Commons Attribution-NoDerivs license
CC  -----
DR  EMBL, AP007255; BAB9553.1; -; Genomic DNA.
KW  Inner membrane; Membrane; Transmembrane; Transport.
SQ  SEQUENCE 458 AA; 48181 MW; 6F13CC647D77468B CRC64;

Query Match          89.5%; Score 34; DB 2; Length 458;
Best Local Similarity 85.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY  1 WMLSAFS 7
    |||||
Db  226 WILSAFS 232

RESULT 17
Q44233 ARTSP PRELIMINARY; PRT; 637 AA.
ID  Q44233 ARTSP PRELIMINARY; PRT; 637 AA.
AC  Q44233;
DT  01-NOV-1996, integrated into UniProtKB/TrEMBL.
DT  01-NOV-1996, sequence version 1.
DT  07-FEB-2006, entry version 20.
DE  Beta-galactosidase.
OS  Arthrobacter sp.
OC  Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC  Micrococcineae; Micrococcaceae; Arthrobacter.
OX  NCBI_TaxId=1667;
RN  [1]
RP  NUCLEOTIDE SEQUENCE.
RX  MEDLINE=95238267; PubMed=7721689;
RA  Gushki L.R., Trimburt D.E., Kasim J.J., Brenshley J.E.;
RT  "Analysis of a novel gene and beta-galactosidase isozyme from a
RT  psychrotrophic Arthrobacter isolate.";
RL  J. Bacteriol. 177:1981-1988(1995).
CC  -----
CC  Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC  Distributed under the Creative Commons Attribution-NoDerivs license
CC  -----
DR  EMBL, U17417; AAA75601.1; -; Genomic DNA.
DR  GO: GO:0009341; C:beta-galactosidase complex; IEA.
DR  GO: GO:0004565; F:beta-galactosidase activity; IEA.
DR  GO: GO:0005975; P:carbohydrate metabolism; IEA.
DR  InterPro: IPR003476; Glyco_hydro_42.
DR  Pfam; PF02449; Glyco_hydro_42; 1.
DR  PIRSF; PIRSF001084; B-galactosidase; 1.
SQ  SEQUENCE 637 AA; 70895 MW; 9816DB4FCC3A728C CRC64;

Query Match          89.5%; Score 34; DB 2; Length 637;
Best Local Similarity 85.7%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY  1 WMLSAFS 7
    |||||
Db  372 WMLSCFS 378

RESULT 18
P88944 HHV8 PRELIMINARY; PRT; 843 AA.
ID  P88944 HHV8 PRELIMINARY; PRT; 843 AA.
AC  P88944;
DT  01-MAY-1997, integrated into UniProtKB/TrEMBL.
DT  01-MAY-1997, sequence version 1.
DT  07-FEB-2006, entry version 23.
DE  ORF 56 (DNA replication protein).
OS  Human herpesvirus 8 (HHV-8) (Kaposi's sarcoma-associated herpesvirus).
OC  Viruses; dsDNA viruses (HHV-8) (Kaposi's sarcoma-associated herpesvirus).
OC  Gammaherpesvirinae; Rhadinovirus.
OX  NCBI_TaxId=37296;

```

```

RN  [1]
RP  NUCLEOTIDE SEQUENCE.
RX  MEDLINE=97094384; PubMed=8939871; DOI=10.1126/science.274.5293.1739;
RA  Moore P.S., Bashoff C., Weiss R.A., Chang Y.;
RT  "Molecular mimicry of human cytokine and cytokine response pathway
RT  genes by KSHV.";
RL  Science 274:1739-1744(1996).
RN  [2]
RP  NUCLEOTIDE SEQUENCE.
RX  MEDLINE=97121480; PubMed=8962146; DOI=10.1073/pnas.93.25.14862;
RA  Russo J.J., Bohenzky R.A., Chien M.-C., Chen J., Yan M., Madgalena D.,
RA  Parry J.P., Peruzzi D., Edelman I.S., Chang Y., Moore P.S.;
RT  "Nucleotide sequence of the Kaposi sarcoma-associated herpesvirus
RT  (HHV8).";
RL  Proc. Natl. Acad. Sci. U.S.A. 93:14862-14867(1996).
RN  [3]
RP  NUCLEOTIDE SEQUENCE.
RA  Russo J.J., Bohenzky R.A., Chien M.-C., Chen J., Yan M., Madgalena D.,
RA  Parry J.P., Peruzzi D., Edelman I.S., Chang Y., Moore P.S.;
RL  Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN  [4]
RP  NUCLEOTIDE SEQUENCE.
RX  MEDLINE=97296220; PubMed=9151804;
RA  Neipel F., Albrecht J.-C., Ensser A., Huang Y.-Q., Li J.J.,
RT  "Cell-homologous genes in the Kaposi's sarcoma-associated rhadinovirus
RT  human herpesvirus 8: determinants of its pathogenicity?";
RL  J. Virol. 71:4187-4192(1997).
RN  [5]
RP  NUCLEOTIDE SEQUENCE.
RA  Neipel F., Albrecht J.-C., Ensser A., Huang Y.-Q., Li J.J.,
RA  Friedman-Kien A.E., Fleckenstein B.;
RT  "The genome of human herpesvirus 8 cloned from Kaposi's sarcoma.";
RL  Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC  -----
CC  Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC  Distributed under the Creative Commons Attribution-NoDerivs license
CC  -----
DR  EMBL, U75698; AAC57138.1; -; Genomic DNA.
DR  EMBL, U93872; AAB62614.1; -; Genomic DNA.
DR  GO: GO:0003896; F:DNA primase activity; IEA.
DR  GO: GO:0006260; P:DNA replication; IEA.
DR  InterPro: IPR004340; UL52 UL70.
DR  Pfam; PF03121; Herpes UL52; 1.
SQ  SEQUENCE 843 AA; 95554 MW; 89248BFCB7DBAB0E CRC64;

Query Match          89.5%; Score 34; DB 2; Length 843;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 WMLSAF 6
    |||||
Db  161 WMLSAF 166

RESULT 19
Q3WYR2_9ACTN PRELIMINARY; PRT; 853 AA.
ID  Q3WYR2_9ACTN PRELIMINARY; PRT; 853 AA.
AC  Q3WYR2;
DT  11-OCT-2005, integrated into UniProtKB/TrEMBL.
DT  11-OCT-2005, sequence version 1.
DT  07-MAR-2006, entry version 6.
DE  Valyl-tRNA synthetase, class 1a.
OS  ORFNames=Exy1DRAFT_0191;
OS  Rubrobacter xylophilus DSM 9941.
OC  Bacteria; Actinobacteria; Rubrobacteridae; Rubrobacterales;
OC  Rubrobacteriaceae; Rubrobacteraceae; Rubrobacter.
OX  NCBI_TaxId=266117;
RN  [1]
RP  NUCLEOTIDE SEQUENCE.
RA  STRAIN=DSM 9941;
RG  US DOE Joint Genome Institute (JGI-PGF);
RA  Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RA  Hammon N., Ierani S., Pitluck S., Richardson P.;

```

```

RT "Sequencing of the draft genome and assembly of Rubrobacter
RT xylanophilus DSM 9941."
RT Submitted (JUL-2005) to the EMBL/Genbank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=DSM 9941;
RA US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Rubrobacter xylanophilus
RT DSM 9941."
RL Submitted (JUL-2005) to the EMBL/Genbank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=DSM 9941;
RA US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RA Hammon N., Istant S., Pritchard S., Richardson P.;
RA Submitted (JUN-2005) to the EMBL/Genbank/DBJ databases.
RL Submitted (JUN-2005) to the EMBL/Genbank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -!- CATALYTIC ACTIVITY: ATP + L-valine + tRNA(Val) = AMP + diphosphate
CC + L-valyl-tRNA(Val).
CC -!- SUBUNIT: Monomer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasm (By similarity).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL: AAB0200094; FAN4753.1; -!- Genomic_DNA.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0016874; F:Ligase activity; IEA.
DR GO: GO:0000166; F:nucleotide binding; IEA.
DR GO: GO:0004832; F:valine-tRNA ligase activity; IEA.
DR GO: GO:0006412; P:protein biosynthesis; IEA.
DR GO: GO:0006438; P:valyl-tRNA aminoacylation; IEA.
DR InterPro: IPR002300; tRNA-synt_1a.
DR InterPro: IPR001412; tRNA-synt_1.
DR InterPro: IPR002303; tRNA-synt_val.
DR Pfam: PF00133; tRNA-synt_1; 1.
DR PRINTS: PR00986; TRNASYNTHAL.
DR TIGRfams: TIGR00422; vals; 1.
DR PROSITE: PS00178; AA_TRNA_LIGASE_1; 1.
DR ATP-binding; Aminoacyl-tRNA synthetase; Ligase; Nucleotide-binding;
DR protein biosynthesis.
SQ SEQUENCE 853 AA; 97415 MW; 6E63EFA9C56B32D CRC64;

Query Match 89.5%; Score 34; DB 2; Length 853;
Best Local Similarity 85.7%; Pred. No. 3.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
DB 607 WILSAFS 613

RESULT 20
Q73RB9 TREDE PRELIMINARY; PRT: 1030 AA.
AC Q73RB9_
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-MAR-2006, entry version 20.
DE Nuclease ShcCD, C subunit, putative.
GN OrderedLocNames=TDE0171; ORFNames=TDE_0171;
OS Treponema denticola.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=158;
RA NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=ATCC 35405 / DSM 14222;
RX PubMed=15064399; DOI=10.1073/pnas.0307639101;
RA Seshadri R., Myers G.S.A., Tetteilin H., Eisen J.A., Heidelberg J.F.,

```

```

RA Dodson R.J., Davidsen T.M., DeBoy R.T., Fouts D.E., Haft D.H.,
RA Selengut J., Ren Q., Brinkac L.M., Madupu R., Kolonay J.F.,
RA Durkin S.A., Daugherty S.C., Shetty J., Shvartsbeyn A.,
RA Gebregeorgis E., Geer K., Tsegaye G., Malek J.A., Ayodeji B.,
RA Shatsman S., McLeod M.P., Smajls D., Howell J.K., Pal S., Amin A.,
RA Vashisth P., McNeill T.Z., Xiang Q., Sodergren E., Baca E.,
RA Weinstein G.M., Norris S.O., Fraser C.M., Paulsen I.T.;
RT "Comparison of the genome of the oral pathogen Treponema denticola
RT with other spirochete genomes."
RL Proc. Natl. Acad. Sci. U.S.A. 101:5646-5651 (2004).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL: AE017226; AAS10668.1; -!- Genomic_DNA.
DR GenomesReviews: AE017226_GR; TDE0171.
DR TIGR: TDE0171; -.
DR BioCyc: TDE0171; TDE0171-MONOMER; -.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0016887; F:ATPase activity; IEA.
DR InterPro: IPR003439; ABC_transp_like.
DR Complete proteome.
SQ SEQUENCE 1030 AA; 117667 MW; EA7480ED35817387 CRC64;

Query Match 89.5%; Score 34; DB 2; Length 1030;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WMLSAF 6
DB 877 WMLSAF 882

RESULT 21
Q4B833 BURVI PRELIMINARY; PRT: 1196 AA.
ID Q4B833_BURVI
AC Q4B833;
DT 13-SEP-2005, integrated into UniProtKB/TrEMBL.
DT 13-SEP-2005, sequence version 1.
DT 07-FEB-2006, entry version 2.
DE Pyruvate decarboxylase; Pyruvate ferredoxin/flavodoxin oxidoreductase.
GN ORFNames=Bcep1808DRPT_0224;
OS Burkholderia vietnamiensis G4.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia; Burkholderia cepacia complex.
OX NCBI_TaxID=269482;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=G4;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RA Hammon N., Istant S., Pritchard S., Richardson P.;
RT "Sequencing of the draft genome and assembly of Burkholderia
RT vietnamiensis G4."
RL Submitted (JUN-2005) to the EMBL/Genbank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=G4;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Burkholderia vietnamiensis
RT G4."
RL Submitted (JUN-2005) to the EMBL/Genbank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=G4;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RA Hammon N., Istant S., Pritchard S., Richardson P.;
RA Submitted (JUN-2005) to the EMBL/Genbank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.

```

```

CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL; AAE0200080; EAM26009.1; -; Genomic DNA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
KW Purvate
SQ SEQUENCE 1196 AA; 129821 MW; 69FF40127361D3F4 CRC64;

Query Match 89.5%; Score 34; DB 2; Length 1196;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WMLSAF 6
DB 1091 WMLSAF 1096

RESULT 22
Q5QTR8_HUMAN PRELIMINARY; PRT; 44 AA.
AC Q5QTR8;
DT 04-JAN-2005, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, entry version 4.
DS Putative P008-2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Aorta;
RA Liu B., Zhao B., Zhang X.Y., Xu Y.Y., Liu Y.Q., Song L., Ye J.,
RA Sheng H., Gao Y., Zhang C.L., Wei Y.J., Zhang J., Song L., Jiang Y.X.,
RA Zhao Z.W., Ding J.F., Liu L.S., Gao R.L., Wu Q.Y., Qiang B.O.,
RA Yuan J.G., Liew C.C., Zhao M.S., Hui R.T.,
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL; AF109190; AAQ13500.1; -; mRNA.
SQ SEQUENCE 44 AA; 5210 MW; 90754272D3B001A4 CRC64;

Query Match 86.8%; Score 33; DB 2; Length 44;
Best Local Similarity 85.7%; Pred. No. 24;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
DB 9 WMLSAFS 15

RESULT 23
Q2UD08 ASPOR PRELIMINARY; PRT; 304 AA.
AC Q2UD08;
DT 24-JAN-2006, integrated into UniProtKB/TrEMBL.
DT 07-MAR-2006, entry version 1.
DS Hydrolyases of the alpha/beta superfamily.
GN ORFNames=AO090012000367;
OS Aspergillus oryzae.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillius.
CX NCBI_TaxID=5062;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=RIB 40;
RC PubMed=16372010; DOI=10.1038/nature04300;
RX

```

```

RA Machida M., Asai K., Sano M., Tanaka T., Kumagai T., Terao G.,
RA Kusunoto K., Arita T., Akita O., Kashiwagi Y., Abe K., Gomi K.,
RA Horiuchi H., Kitamoto K., Kobayashi T., Takeuchi M., Denning D.W.,
RA Horikawa J.E., Nierman W.C., Yu J., Archer D.B., Bennett J.W.,
RA Bhattacharya D., Cleveland T.E., Fedorova N.D., Gotch O., Horikawa H.,
RA Hosoyama A., Ichinomiya M., Igarashi R., Iwashita K., Juvvadi P.R.,
RA Kato M., Kato Y., Kin T., Kokubun A., Maeda H., Maeyama N.,
RA Maruyama J., Nagasaki H., Nakajima T., Oda K., Okada K., Paulsen I.,
RA Sakamoto K., Sawano T., Takahashi M., Takase K., Terabayashi Y.,
RA Wortman J.R., Yamada O., Yamagata Y., Anazawa H., Hata A., Isono K.,
RA Komori T., Koyama Y., Minetoki T., Suharnan S., Tanaka Y., Isono K.,
RA Kuhnara S., Ogasawara N., Kikuchi H.;
RT "Genome sequencing and analysis of Aspergillus oryzae.";
RL Nature 438:1157-1161(2005).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL; AP007161; BAB60557.1; -; Genomic DNA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
KW Hydrolase.
SQ SEQUENCE 304 AA; 33626 MW; A3D3EA35882B9D6A CRC64;

Query Match 86.8%; Score 33; DB 2; Length 304;
Best Local Similarity 85.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
DB 168 WMLSLFS 174

RESULT 24
PRG4_HUMAN STANDARD; PRT; 1404 AA.
ID PRG4_HUMAN
AC Q92954; O6DNC4; O6DNC5; O6ZM25; Q9BX49;
DT 06-DEC-2005, integrated into UniProtKB/Swiss-Prot.
DT 06-DEC-2005, entry version 2.
DT 07-MAR-2006, entry version 38.
DE Proteoglycan-4 precursor (ubridicin) (Megakaryocyte stimulating factor)
DE (Superficial zone proteoglycan) [Contains: Proteoglycan-4 C-terminal
DE part].
GN Name=PRG4; Synonyms=MSF, SZP;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM A), AND VARIANT TRP-180.
RA Turner K.J., Fitz L.J., Temple P., Jacobs K., Larson D., Leary A.C.,
RA Kelleher K., Giannotti J., Calvetti J., Fitzgerald M., Kriz M.J.,
RA Ferenz C., Grobholz J., Fraser H., Bean K., Norton C.R., Gesner T.,
RA Bhatia S., Kriz R., Hewick R., Clark S.C.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM E).
RC TISSUE=Synovial cell;
RC PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura T., Makita H.,
RA Sekine M., Odayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto K., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahari K.,
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
RA Niimiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hiraoaka S., Chiba Y., Ishida S.,
RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Houta T., Kuano Y.,
RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O., Nomura Y.,
RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
RA Muraashino K., Yuki H., Oshima A., Sasaki N., Aotaka S.,

```

RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
 RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
 RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
 RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
 RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujitara T.,
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hiroo M., Oimoto Y.,
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Ohtani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
 RA Matsunura K., Nakajima Y., Mizuno T., Moriga M., Sasaki M.,
 RA Togaishi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
 RA Minushima-Sugano J., Sato T., Shirai Y., Takahashi Y., Nakagawa K.,
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuno Y., Yamashita R.,
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.,
 RT "Complete sequencing and characterization of 21,243 full-length human
 RT cDNAs.";
 RL Nat. Genet. 36:40-45(2004).
 RN [3]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RG Human chromosome 1 international sequencing consortium;
 RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP NUCLEOTIDE SEQUENCE [MRNA] OF 32-193 AND 1148-1398.
 RA Jones A.R., Hughes C.E., Flannery C.R., Caterson B.,
 RT "Cloning and production of recombinant PRG4/cartilage superficial zone
 RT proteoglycan (SZP) N- and C-terminal domains.";
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP NUCLEOTIDE SEQUENCE [MRNA] OF 39-90; 158-258 AND 1209-1295, AND
 RP IDENTIFICATION OF ISOFORM C.
 RX MEDLINE=99120896; PubMed=9920774; DOI=10.1006/dbrc.1998.0104.
 RA Flannery C.R., Hughes C.E., Schumacher B.L., Tudor D., Aydelotte M.B.,
 RA Kuettner K.E., Caterson B.,
 RT "Articular cartilage superficial zone protein (SZP) is homologous to
 RT megakaryocyte stimulating factor precursor and is a multifunctional
 RT proteoglycan with potential growth-promoting, cytoprotective, and
 RT lubricating properties in cartilage metabolism.";
 RL Biochem. Biophys. Res. Commun. 254:535-541(1999).
 RN [6]
 RP NUCLEOTIDE SEQUENCE [MRNA] OF 41-376 (ISOFORM F), FUNCTION, AND
 RP SUBCELLULAR LOCATION.
 RC Tissue=Fetal liver;
 RX PubMed=14976050; DOI=10.1182/blood-2003-06-1825.
 RA Liu Y.-J., Lu S.H., Xu B., Yang R.C., Ren Q., Liu B., Li B., Lu M.,
 RA Yan F.Y., Han Z.B., Han Z.C.,
 RT "Hemangiopoietin, a novel human growth factor for the primitive cells
 RT of both hematopoietic and endothelial cell lineages.";
 RL Blood 103:4449-4456(2004).
 RN [7]
 RP PURIFICATION.
 RP Tissue=Urine;
 RA Turner K.J., Fitz L.J., Temple P., Jacobs K., Larson D., Leary A.C.,
 RA Kuettner K., Giannotti J., Calvetti J., Fitzgerald M., Kriz M.-J.,
 RA Perenz C., Grobholz J., Fraser H., Bean K., Norton C.R., Gesser T.,
 RA Bhargava S., Kriz R., Hewick R., Clark S.C.,
 RT "Purification, biochemical characterization, and cloning of a novel
 RT megakaryocyte stimulating factor that has megakaryocyte colony
 RT stimulating activity.";
 RL Blood 78:279A-279A(1991).
 RN [8]
 RP GENE STRUCTURE.
 RA Weisberg D.M., Fitz L.J., Temple P., Giannotti J., Murtha P.,
 RA Fitzgerald M., Scaltreto H., Kuettner K., Preissner K., Kriz R.,
 RA Jacobs K., Turner K.,
 RT "A comparison of vitronectin and megakaryocyte stimulating factor.";
 RL (In) Preissner K.T., Rosenblatt S., Kost C., Wegerhoff J.,
 MOESER D.F. (eds.),
 RL Biology of vitronectins and their receptors, pp.45-52, Elsevier
 RL Science Publishers B.V., Amsterdam (1993).
 RN [9]
 RP GLYCOSYLATION.
 RX MEDLINE=94241694; PubMed=8185311; DOI=10.1006/abbi.1994.1219;
 RA Schumacher B.L., Block J.A., Schmid T.M., Aydelotte M.B.,

RA Kuettner K.E.;
 RT "A novel proteoglycan synthesized and secreted by chondrocytes of the
 RT superficial zone of articular cartilage.";
 RL Arch. Biochem. Biophys. 311:144-152(1994).
 RN [10]
 RP TISSUE SPECIFICITY, AND INVOLVEMENT IN CACP.
 RX PubMed=10545550; DOI=10.1038/154496;
 RA Marcelino J., Caplen J.D., Suwalit W.M., Gutierrez O.M., Schwartz S.,
 RA Robbins C., Sood R., Makalowska I., Baxevanis A., Johnstone B.,
 RA Laxer R.M., Zemel L., Kim C.A., Herd J.K., Ihle J., Williams C.,
 RA Johnson M., Ramon V., Alonso L.G., Brunoni D., Gerstein A.,
 RA Papadopoulos N., Bahabri S.A., Trent J.M., Marman M.L.,
 RT "CACP, encoding a secreted proteoglycan, is mutated in campodactyl-
 RT arthropathy-coxa vara-pericarditis syndrome.";
 RL Nat. Genet. 23:319-322(1999).
 RN [11]
 RP TISSUE SPECIFICITY, AND IDENTIFICATION OF ISOFORMS B; C AND D.
 RX MEDLINE=20573856; PubMed=11124536;
 RA Ikegawa S., Sano M., Koshizuka Y., Nakamura Y.,
 RT "Isolation, characterization and mapping of the mouse and human PRG4
 RT (proteoglycan 4) genes.";
 RL Cytogenet. Cell Genet. 90:291-297(2000).
 RN [12]
 RP IDENTIFICATION BY MASS SPECTROMETRY, FUNCTION, AND GLYCOSYLATION.
 RX PubMed=10743795;
 RA Day G.D., Britz D.E., Cha C.-J.,
 RT "Lubricin is a product of megakaryocyte stimulating factor gene
 RT expression by human synovial fibroblasts.";
 RL J. Rheumatol. 27:594-600(2000).
 RN [13]
 RP CARBOHYDRATE-LINKAGE SITE ASN-1159.
 RX PubMed=16335952; DOI=10.1021/pr0502065;
 RA Liu T., Qian W.-J., Gritsenko M.A., Camp D.G. II, Monroe M.E.,
 RA Moore R.J., Smith R.D.,
 RT "Human plasma N-glycoproteome analysis by immunofluorescence
 RT hydrazone chemistry, and mass spectrometry.";
 RL J. Proteome Res. 4:2070-2080(2005).
 CC -I- FUNCTION: Plays a role for boundary lubrication within
 CC articulating joints. Prevents protein deposition onto cartilage
 CC from synovial fluid, control adhesion-dependent synovial growth,
 CC and inhibit the adhesion of synovial cells to the cartilage
 CC surface.
 CC -I- FUNCTION: Isoform F play a role as a growth factor acting on the
 CC primitive cells of both hematopoietic and endothelial cell
 CC lineages.
 CC -I- SUBUNIT: Homodimer; disulfide-linked (By similarity).
 CC -I- SUBCELLULAR LOCATION: Secreted protein.
 CC -I- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=6;
 CC Name=A;
 CC IsoId=Q92954-1; Sequence=Displayed;
 CC Name=B;
 CC IsoId=Q92954-2; Sequence=VSP_016467;
 CC Name=C;
 CC IsoId=Q92954-3; Sequence=VSP_016468;
 CC Name=D;
 CC IsoId=Q92954-4; Sequence=VSP_016467, VSP_016468;
 CC Name=E;
 CC IsoId=Q92954-5; Sequence=VSP_016467, VSP_016470;
 CC Note=No experimental confirmation available;
 CC Name=F; Synonyms=Hemangiopoietin, HAP0;
 CC IsoId=Q92954-6; Sequence=VSP_016469;
 CC -I- TISSUE SPECIFICITY: Highly expressed in synovial tissue, cartilage
 CC and liver, and weakly in heart and lung. Isoform B is expressed in
 CC kidney, lung, liver, heart and brain. Isoform C and isoform D are
 CC widely expressed.
 CC -I- PTM: N-glycosylated.
 CC -I- PTM: O-glycosylated; contains glycosaminoglycan chondroitin
 CC sulfate and keratan sulfate.
 CC -I- PTM: The disulfide bond between Cys-1146 and Cys-1403 is essential
 CC for protein cleavage (By similarity).
 CC -I- DISEASE: Defects in PRG4 are the cause of campodactyl-
 CC arthropathy-coxa vara-pericarditis syndrome (CACP) [MIM:208250];

```

CC also called Jacobs syndrome. CACP is an autosomal recessive
CC disorder. Individuals with CACP have normal appearing joints at
CC birth but with advancing age develop joint failure associated with
CC noninflammatory synovioocyte hyperplasia and subintimal fibrosis of
CC the synovial capsule.
CC -1- MISCELLANEOUS: Different forms of various molecular weight have
CC been observed. Such forms are possibly due to different levels of
CC glycosylation and protein cleavage (By similarity).
CC -1- SIMILARITY: Contains 2 hemopexin-like domains.
CC -1- SIMILARITY: Contains 2 SMR (somatomedin B) domains.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC
DR EMBL; U70136; AAB09089.1; -; mRNA.

Query Match      86.8%; Score 33; DB 1; Length 1404;
Best Local Similarity 85.7%; Pred. No. 8.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 WMLSAFS 7
Db 1171 WMLSPFS 1177

RESULT 25
Q46T11_RALEJ PRELIMINARY; PRT; 74 AA.
AC Q46T11;
DT 13-SEP-2005, integrated into UniProtKB/TrEMBL.
DT 13-SEP-2005, sequence version 1.
DT 07-FEB-2006, entry version 4.
DE Hypothetical protein.
GN OrderedLocNames=Reut_B4097;
OS Ralstonia eutropha (strain JMP134) (Alcaligenes eutrophus).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Cupriavidus.
OX NCBI_TaxID=264198;
RN NM
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler J.C., Glavina T.,
RA Hammon N., Istant S., Plick S., Goldsman E., Martinez M.,
RA Schmutz J., Larimer F., Land M., Lykidis A., Richardson P.;
RT "Complete sequence of chromosome 2 of Ralstonia eutropha JMP134.";
RL Submitted (Aug-2005) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC
DR EMBL; CP000091; AA263453.1; -; Genomic DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 74 AA; 8424 MW; 39E3984EB72662AE CRC64;

Query Match      84.2%; Score 32; DB 2; Length 74;
Best Local Similarity 71.4%; Pred. No. 67;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WMLSAFS 7
Db 23 WMLAFA 29

RESULT 26
Q4S8Z0_TETNG PRELIMINARY; PRT; 117 AA.
AC Q4S8Z0;
DT 19-JUL-2005, integrated into UniProtKB/TrEMBL.
DT 19-JUL-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Chromosome 7 SCAR14703, whole genome shotgun sequence.
GN ORNames=GSTENG00023129001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN NM
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15496914; DOI=10.1038/nature03025;
RA Jallion O., Aury J.-M., Brunet F., Petit J.-L., Strange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicoud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Da Silva C., Salanoubat M., Levy M., Boudet N., Castellan S.,
RA Anhouard V., Jabin C., Castelli V., Katinka M., Vacherie B.,
RA Bismont C., Skalli Z., Cactolico L., Poullain J., De Bernardis V.,
RA Crnaud C., Duprat S., Brotier P., Couanceau J.-P., Gouzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volt J.-N., Guigo R., Zody M.C., Westrov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Lauder V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissenbach J., Roest Crolius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN NM
RP NUCLEOTIDE SEQUENCE.
RG Genoscope, Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC
DR EMBL; CAE01014703; CAC02892.1; -; Genomic DNA.
SQ SEQUENCE 117 AA; 12900 MW; 8D07018118B6F483 CRC64;

Query Match      84.2%; Score 32; DB 2; Length 117;
Best Local Similarity 71.4%; Pred. No. 1.1e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WMLSAFS 7
Db 30 WMLSAFS 36

RESULT 27
Q43Q96_SOLUS PRELIMINARY; PRT; 129 AA.
AC Q43Q96;
DT 13-SEP-2005, integrated into UniProtKB/TrEMBL.
DT 13-SEP-2005, sequence version 1.
DT 07-FEB-2006, entry version 2.
DE Transport-associated precursor.
GN ORNames=AcidDRAFT_1152;
OS Solibacter usitatus Ellin6076
OC Bacteria; Acidobacteria; Solibacteres; Solibacterales;
OC Solibacteraceae; Solibacter.
OX NCBI_TaxID=234267;
RN NM
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Ellin6076;
RG US DOE Joint Genome Institute (JGI-RGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RA Hammon N., Istant S., Plick S., Goldsman E., Martinez M.,
RT "Sequencing of the draft genome and assembly of Solibacter usitatus
RT Ellin6076.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN NM
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Ellin6076;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Solibacter usitatus.";

```

```

RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL; AAI0100063; EAM54975.1; -; Genomic_DNA.
DR InterPro; IPR007055; TAD.
DR Pfam; PF04972; BON; 1.
DR PROSITE; PS50914; BON; 1.
DR Signal.
FT SIGNAL
SQ SEQUENCE 129 AA; 13518 MW; A7290632DFF510E4 CRC64;

Query Match 84.2%; Score 32; DB 2; Length 129;
Best Local Similarity 71.4%; Pred. No. 1.2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 1 WMLSAFS 7
Db 2 WMLNAFT 8

RESULT 28
O9WKJ3 PORGI PRELIMINARY; PRT; 159 AA.
ID O9WKJ3 PORGI
AC O9WKJ3
DT 01-NOV-1999, integrated into UniProtKB/TrEMBL.
DT 01-NOV-1999, sequence version 1.
DT 07-FEB-2006, entry version 22.
DE Dps.
GN Name=dps;
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteri; Bacteroidetes; Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RX MEDLINE=22483669; PubMed=12595429;
RA DOI=10.1128/IAI.71.3.1170-1178.2003;
RA Ueshima J., Shoji M., Ratnayake D.B., Abe K., Yoshida S., Yamamoto K.,
RA Nakayama K.;
RT "Purification, gene cloning, gene expression, and mutants of Dps from
RT the obligate anaerobe Porphyromonas gingivalis.";
RL Infect. Immun. 71:1170-1178(2003).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL; AB025779; BAA76886.1; -; Genomic_DNA.
DR HSSP; Q8RP01; IUG.
DR GO; GO:0008199; F:ferriic iron binding; IEA.
DR GO; GO:0006879; P:iron ion homeostasis; IEA.
DR GO; GO:0006950; P:response to stress; IEA.
DR InterPro; IPR002177; Dps DNA bd.
DR InterPro; IPR008331; Ferritin_Dps.
DR InterPro; IPR012347; Ferritin_rel.
DR Pfam; PF00210; Ferritin; 1.
DR PRINTS; PR01346; HELNAPAPROT.
DR PRODOM; PD149803; DPS; 1.
DR PROSITE; PS00818; DPS_1; 1.
DR PROSITE; PS00819; DPS_2; 1.
SQ SEQUENCE 159 AA; 17793 MW; 60889B904D33184C CRC64;

Query Match 84.2%; Score 32; DB 2; Length 159;
Best Local Similarity 71.4%; Pred. No. 1.5e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 1 WMLSAFS 7
Db 2 WMLNAFT 8

```

```

Db 151 WMLSAVA 157

RESULT 29
O7MXS1 PORGI PRELIMINARY; PRT; 159 AA.
ID O7MXS1 PORGI
AC O7MXS1
DT 15-DEC-2003, integrated into UniProtKB/TrEMBL.
DT 15-DEC-2003, sequence version 1.
DT 07-FEB-2006, entry version 14.
DE Dps family protein.
GN OrderedOcuscNames=PG0090; ORFNames=PG_0090;
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RX MEDLINE=22829867; PubMed=12949112;
RA DOI=10.1128/JB.185.18.5591-5601.2003;
RA Nelson K.E., Fleischmann R.D., DeBoy R.T., Paulsen I.T., Fouts D.E.,
RA Eisen J.A., Daugherty S.C., Dodson R.J., Durkin A.S., Gwinn M.L.,
RA Haft D.H., Kolonay J.F., Nelson W.C., Mason T.M., Tallon L., Gray J.,
RA Granger D., Tettelin H., Dong H., Galvin J.L., Duncan W.C.,
RA Dewhirst F.E., Fraser C.M.;
RT "Complete genome sequence of the oral pathogenic bacterium
RT Porphyromonas gingivalis strain W83.";
RL J. Bacteriol. 185:5591-5601(2003).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL; AE015924; AA065337.1; -; Genomic_DNA.
DR HSSP; Q8RP01; IUG.
DR TIGR; PG0090; -.
DR Biocyc; PGIN242619; PG0090-MONOMER; -.
DR GO; GO:0008199; F:ferriic iron binding; IEA.
DR GO; GO:0006879; P:iron ion homeostasis; IEA.
DR GO; GO:0006950; P:response to stress; IEA.
DR InterPro; IPR002177; Dps DNA bd.
DR InterPro; IPR008331; Ferritin_Dps.
DR InterPro; IPR012347; Ferritin_rel.
DR Pfam; PF00210; Ferritin; 1.
DR PRINTS; PR01346; HELNAPAPROT.
DR PRODOM; PD149803; DPS; 1.
DR PROSITE; PS00818; DPS_1; 1.
DR PROSITE; PS00819; DPS_2; 1.
KW Complete proteome.
SQ SEQUENCE 159 AA; 17865 MW; 60889B904D331AC6 CRC64;

Query Match 84.2%; Score 32; DB 2; Length 159;
Best Local Similarity 71.4%; Pred. No. 1.5e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 1 WMLSAFS 7
Db 151 WMLSAVA 157

RESULT 30
O84S36 ORYSA PRELIMINARY; PRT; 167 AA.
ID O84S36 ORYSA
AC O84S36
DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2003, sequence version 1.
DT 07-FEB-2006, entry version 9.
DE Hypothetical protein OJ9990_A01.131 (Hypothetical protein
DE OJ1449_C01.14).
GN Name=OJ9990_A01.131; Synonyms=OJ1449_C01.14;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Bp clade;

```

```

OC Euhartioideae; Oryzeae; Oryza.
RX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Katayose Y.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 8, BAC
  clone:OJ9990 A01."
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Katayose Y.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 8, BAC
  clone:OJ1449 C01."
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL; AP005847; BAC99919.1; -; Genomic DNA.
DR EMBL; AP005605; BAC99823.1; -; Genomic DNA.
DR Gramene; O84536; -.
KW Hypothetical protein.
SQ SEQUENCE 167 AA; 18427 MW; 367D8ABC862C1BBB CRC64;

Query Match      84.2%; Score 32; DB 2; Length 167;
Best Local Similarity 71.4%; Pred. No. 1.6e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
   |||:|
   58 WMLAFA 64

RESULT 31
Q4SFB2.TETNG PRELIMINARY; PRT; 182 AA.
ID Q4SFB2.TETNG
AC Q4SFB2;
DT 19-JUL-2005, integrated into UniProtKB/TrEMBL.
DT 19-JUL-2005, sequence version 1.
DE 07-FEB-2006, entry version 3.
GN Chromosome 6 SCAP14605, whole genome shotgun sequence.
OS Oryza sativa (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15496914; DOI=10.1038/nature03025;
RA Jallion O., Aury J.-M., Brunet F., Petit J.-L., Stange-Thomann N.,
RA Mauceli E., Bounneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dastilva C., Salanoubat M., Levy M., Boudet N., Castellan S.,
RA Anthonard V., Jodin C., Castelli V., Kacinka M., Vacherie B.,
RA Blomont C., Skalli Z., Catolico L., Poulain J., De Bernardis V.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.-N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Lander V., Schachter V., Queller F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Croillius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
  the early vertebrate proto-karyotype."
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope, Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.

```

```

CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL; CAE01014605; CA000670.1; -; Genomic DNA.
SQ SEQUENCE 182 AA; 20009 MW; 0869B43C89CB7475 CRC64;

Query Match      84.2%; Score 32; DB 2; Length 182;
Best Local Similarity 75.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
   |||:|
   71 WMLAFA 77

RESULT 32
PUT2.FUGRU STANDARD; PRT; 187 AA.
ID PUT2.FUGRU
AC 07698;
DT 10-MAY-2002, integrated into UniProtKB/Swiss-Prot.
DT 01-AUG-1998, sequence version 1.
DT 07-FEB-2006, entry version 28.
DE Putative protein 2 (PUT2) (Fragment).
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Takifugu.
OX NCBI_TaxID=31033;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RA Riboldi Tumiciliffe G.R., Platzer M., Nyakatura G., Elgar G.S.,
RA Brenner S., Rosenthal A.;
RT "Analysis of the genomic loci of Fugu rubripes homologs of the human
  disease genes LiCAM, G6PD and P55."
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Membrane; multi-pass membrane protein
  (Potential).
CC -!- SIMILARITY: Belongs to the TMEM9 family.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL; AF026198; AAC15584.1; -; Genomic DNA.
DR PIR; T30536; T30536.
DR Ensemble; SINFU00000152263; Fugu rubripes.
DR GO; GO:0005770; C:late endosome; ISS.
DR GO; GO:0005764; C:lysosome; ISS.
DR InterPro; IPR004153; CXCXC_repeat.
DR InterPro; IPR008853; TMEM9.
DR PANTHER; PTHR13064; TMEM9; 1.
DR Pfam; PF03128; CXCXC; 1.
DR Pfam; PF05434; TMEM9; 1.
KW Hypothetical protein; Membrane; Transmembrane.
FT CHAIN 1 >187 Putative protein 2.
FT TRANSMEM 10 27 /FTid=PRO_0000221039.
FT FT TRANSMEM 99 121 Potential.
FT NON TER 187 187 Potential.
SQ SEQUENCE 187 AA; 21460 MW; 51DD101FB466DD3B CRC64;

Query Match      84.2%; Score 32; DB 1; Length 187;
Best Local Similarity 71.4%; Pred. No. 1.7e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
   |||:|
   9 WMLAFA 15

RESULT 33
Q6CXD3_KLULA

```

```

ID O6CKD3_KLULA PRELIMINARY; PRT; 226 AA.
AC O6CKD3;
DT 16-AUG-2004, integrated into UniProtKB/TrEMBL.
DT 16-AUG-2004, sequence version 1.
DT 07-FEB-2006, entry version 12.
DE Similar to sp|P38869 Saccharomyces cerevisiae YHR181w singleton.
GN OrderedLocustNames=KLUA0A09295g;
OS Kluyveromyces lactis (Yeast).
OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
OX NCBI_TaxID=26985;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=CBS 2359 / IFO 1267 / NRRL Y-1140 / WM37;
RX PubMed=15229592; DOI=10.1038/nature02579;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neveglies C., Talla E.,
RA Goffard N., Prangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Batnay S., Blanchin S., Beckerich J.-M., Beyne E., Bleykasten C.,
RA Boistrane A., Boyer J., Catolico L., Confanier F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantreay F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicand J.-M., Nikolaki M., Ozas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.-F., Straub M.-L., Suleau A.,
RA Swennen D., Tekata F., Wesolowski-Louvel M., Westhof E., With B.,
RA Zentou-Meyer M., Zivanovic Y., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gallardin C., Weissenbach J.,
RA Wincker P., Souciet J.-L.;
RL "Genome evolution in yeasts."
RL Nature 430:35-44(2004).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL: CR382121; CAH02994.1; -; Genomic_DNA.
DR InterPro: IPR007277; DUF396.
DR PANTHER: PTHR13144; DUF396; 1.
DR Pfam: PF04148; DUF396; 1.
DR Complete proteome.
DR KW Complete proteome.
SQ SEQUENCE 226 AA; 26133 MW; D45342F19A9CCEID CRC64;

Query Match 84.2%; Score 32; DB 2; Length 226;
Best Local Similarity 71.4%; Pred. No. 2.1e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 WMLSAFS 7
   |||||
Db 66 WMLTAFS 72

RESULT 34
O6ZRV1_HUMAN PRELIMINARY; PRT; 243 AA.
AC O6ZRV1;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 21-FEB-2006, entry version 13.
DE CDNA FJ446073 file, clone TEST12001364, highly similar to Homo sapiens
DE lactate dehydrogenase A-like (LDHL).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Testis;
RA Niimiya K., Wagatsuma M., Kanda K., Kondo H., Yokoi T., Kodaira H.,
RA Furuya T., Takahashi M., Kikkawa E., Omura Y., Abe K., Kamihara K.,
RA Katsuta N., Sato K., Tanikawa M., Yamazaki M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,

```

```

RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Murakawa K.,
RA Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahari K., Masuno Y., Nagai K., Isogai T.;
RT "NDO human cDNA sequencing project."
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL: AK127961; BAC87208.1; -; mRNA.
DR HSSP: P16115; 1A5Z.
DR GO: GO:0004459; P:L-lactate dehydrogenase activity; IEA.
DR GO: GO:0016491; F:oxidoreductase activity; IEA.
DR GO: GO:0006096; P:glycolysis; IEA.
DR GO: GO:0006100; P:tricarboxylic acid cycle intermediate metab. .; IEA.
DR InterPro: IPR001557; L_LDH_NDH.
DR InterPro: IPR011304; L_LDH_NDH.
DR InterPro: IPR001236; lact_mal_DH.
DR InterPro: IPR002020; NAD_ES.
DR InterPro: IPR000594; ThF_NAD_FAD_bd.
DR PANTHER: PTHR11540; Ldh; 1.
DR Pfam: PF02866; Ldh_1_C; 1.
DR Pfam: PF00056; Ldh_1_N; 1.
DR PRINTS: PR00086; LLDHDSGNSE.
DR PROSITE: PS00064; L_LDH; UNKNOWN 1.
SQ SEQUENCE 243 AA; 26419 MW; 98D3B92406208F25 CRC64;

Query Match 84.2%; Score 32; DB 2; Length 243;
Best Local Similarity 85.7%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 WMLSAFS 7
   |||||
Db 187 WMLTAFS 193

RESULT 35
P72049_MYCTU PRELIMINARY; PRT; 280 AA.
AC P72049; Q7D4V5;
DT 01-FEB-1997, integrated into UniProtKB/TrEMBL.
DT 01-FEB-1997, sequence version 1.
DT 07-FEB-2006, entry version 31.
DE PROBABLE O-ANTIGEN/LIPOPOLYSACCHARIDE TRANSPORT INTEGRAL MEMBRANE
DE PROTEIN ABC TRANSPORTER REPD (O-antigen export system, permease
DE protein).
GN Name=rfbd; OrderedLocustNames=WT3892, RV3783;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OX NCBI_TaxID=1773;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=H37RV;
DE MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
RA Harris D.E., Gordon S.V., Eigemeier K., Gae S., Barry C.E. III,
RA Tekala F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Devlin K., Fellwell T., Gentles S., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Krogh A., McLaren J., Moule S.,
RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RA Rogers J.E., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=CDC 1551 / Oshkosh;
DE MEDLINE=22206494; PubMed=12218036;
RX DOI=10.1128/JB.184.19.5479-5490.2002;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,

```


RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
RA Hickey E.K., Kolonay J.F., Nelson W.C., Umayan L.A., Ermolaeva M.D.,
RA Salzberg S.L., Delcher A., Ustebay T.R., Weidman J.F., Khouli H.M.,
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.,
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL J. Bacteriol. 184:5479-5490(2002).
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL: BX842584; CAB02462.1; -; Genomic DNA.
DR EMBL: AE000516; AK48257.1; -; Genomic DNA.
DR PIR: C70696; C70696.
DR TIGR: MT3892; -.
DR Tuberculist; RV3783; -.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR000412; ABC_2.
DR Pfam: PF01061; ABC2_membrane; 1.
KM Complete proteome.
SQ SEQUENCE 280 AA; 32336 MW; D8793899C05A7EC3 CRC64;

Query Match 84.2%; Score 32; DB 2; Length 280;
Best Local Similarity 71.4%; Pred. No. 2.6e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WMLSAFS 7
Db 262 WMLAFA 268

RESULT 36
O7TVN9_MYCBO PRELIMINARY; PRT; 280 AA.
AC O7TVN9;
DT 01-OCT-2003, integrated into UniProtKB/TrEMBL.
DT 01-OCT-2003, sequence version 1.
DT 07-FEB-2006, entry version 11.
DE PROBABLE O-ANTIGEN/LIPOPOLYSACCHARIDE TRANSPORT INTEGRAL MEMBRANE
DE PROTEIN ABC TRANSPORTER RFBP.
GN Name=rfbD; OrderedLocustNames=MB3812;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1765;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garnier T., Eigmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Piryor M., Dutroy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating P.R.,
RA Parkhill J., Barrett B.G., Cole S.T., Gordon S.V., Hewinson R.G.,
RT "The complete genome sequence of Mycobacterium bovis."
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL: BX248347; CAD95998.1; -; Genomic DNA.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR000412; ABC_2.
DR Pfam: PF01061; ABC2_membrane; 1.
KM Complete proteome.
SQ SEQUENCE 280 AA; 32336 MW; D8793899C05A7EC3 CRC64;

Query Match 84.2%; Score 32; DB 2; Length 280;
Best Local Similarity 71.4%; Pred. No. 2.6e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WMLSAFS 7
Db 262 WMLAFA 268

RESULT 37
MRGRG_MOUSE STANDARD; PRT; 289 AA.
ID MRGRG_MOUSE
AC G912B5; Q711N2;
DT 31-AUG-2004, integrated into UniProtKB/Swiss-Prot.
DT 01-DEC-2001, sequence version 1.
DT 07-FEB-2006, entry version 22.
DE Mas-related G-protein coupled receptor member G (Evolutionary
DE breakpoint transcript 2 protein).
GN Name=Mrgprg; Synonyms=Ebrt2, Mrgs;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RC STRAIN=C57BL/6;
RX MEDLINE=21435808; PubMed=11551509; DOI=10.1016/S0092-8674(01)00483-4;
RA Dong X., Han S.-K., Zylka M.J., Simon M.I., Anderson D.J.,
RT "A diverse family of GPCRs expressed in specific subsets of
RT nociceptive sensory neurons."
RL Cell 106:619-632(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE [MRNA] OF 1-156.
RC STRAIN=C57BL/6;
RA Engemann S., Stroedicke M., Meguro M., Franck O., Kalscheuer V.,
RA Oshimura M., Walter J.,
RT "The BWS region contains an evolutionary breakpoint region."
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Orphan receptor. May regulate nociceptor function and/or
CC development, including the sensation or modulation of pain.
CC -1- SUBCELLULAR LOCATION: Membrane; multi-pass membrane protein.
CC -1- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
CC Mas subfamily.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL: AY042212; AK91803.1; -; Genomic DNA.
DR EMBL: AJ313465; CAC86130.1; -; mRNA.
DR EMBL: ENSMUSG0000050276; Mus musculus.
DR MGI: MGI:3033145; Mrgprg.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR PIR: PIR00001; 7tm.1; 1.
DR PRINTS: PR00237; GPCRHOOPS.
DR PROSITE: PS00237; G-PROTEIN_RECP_F1_1; FALSE NEG.
DR PROSITE: PS50262; G-PROTEIN_RECP_F1_2; FALSE NEG.
KM G-protein coupled receptor; Membrane; Receptor; Transducer;
KM Transmembrane.
FT CHAIN 1 289
FT member G.
FT /Ftld=PRO_0000069768.
FT Extracellular (Potential).
FT TOPO_DOM 1 13
FT TRANSMEM 14 34
FT TOPO_DOM 35 52
FT TRANSMEM 53 73
FT TOPO_DOM 74 78
FT TRANSMEM 79 99
FT TOPO_DOM 100 120
FT TRANSMEM 121 141
FT TOPO_DOM 142 164
FT TRANSMEM 165 185
FT Mas-related G-protein coupled receptor
FT Extracellular (Potential).
FT 5 (Potential).

FT TOPO_DOM 186 199 Cytoplasmic (Potential).
 FT TRANSMEM 200 220 6 (Potential).
 FT TOPO_DOM 221 222 Extracellular (Potential).
 FT TRANSMEM 223 243 7 (Potential).
 FT TOPO_DOM 244 289 Cytoplasmic (Potential).
 FT CONFLICT 122 122 T->I (in Ref.2).
 SQ SEQUENCE 289 AA; 32028 MW; 82BD4BEA02E0864C CRC64;

Query Match 84.2%; Score 32; DB 1; Length 289;
 Best Local Similarity 71.4%; Pred. No. 2.7e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 WMLSAFS 7
 Db 94 WMLAIFS 100

RESULT 38
 Q498A2_MOUSE PRELIMINARY; PRT; 289 AA.
 ID Q498A2_MOUSE
 AC Q498A2;
 DT 13-SEP-2005, integrated into UniProtKB/TrEMBL.
 DT 13-SEP-2005, sequence version 1.
 DT 21-FEB-2006, entry version 7.
 DE MAS-related GPR, member G.
 GN Name=Mrgpr3;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Placenta;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Scheefel C.F., Bat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.D., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kerteman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Placenta;
 RG NIH MGC Project;
 RL Submitted (Aug-2005) to the EMBL/GenBank/DBJ databases.
 CC -1- SOURCE/LOCATOR: Membrane; multi-pass membrane protein (By
 similarity).
 CC Copyrighted under the Creative Commons Attribution-NonDerivs License
 CC Distributed under the Creative Commons Attribution-NonDerivs License
 CC EMBL, EBI00302; AAI00303.1; -, mRNA.
 DR MGI; MGI:3033145; Mrgpr3.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0004872; P:receptor activity; IEA.
 DR GO; GO:0001584; P:rhodopsin-like receptor activity; IEA.
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.

DR GO; GO:0007165; P:signal transduction; IEA.
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCR_RHODOPSN.
 KW G-protein coupled receptor; Membrane; Receptor; Transducer;
 KW Transmembrane.
 SQ SEQUENCE 289 AA; 32028 MW; 82BD4BEA02E0864C CRC64;

Query Match 84.2%; Score 32; DB 2; Length 289;
 Best Local Similarity 71.4%; Pred. No. 2.7e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 WMLSAFS 7
 Db 94 WMLAIFS 100

RESULT 39
 Q3M9S1_ANAVY PRELIMINARY; PRT; 333 AA.
 ID Q3M9S1_ANAVY
 AC Q3M9S1;
 DT 25-OCT-2005, integrated into UniProtKB/TrEMBL.
 DT 25-OCT-2005, sequence version 1.
 DT 07-FEB-2006, entry version 3.
 DE Hypothetical protein precursor.
 GN ORFNames=Ava_2652;
 OS Anabaena variabilis (strain ATCC 29413).
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Anabaena.
 NCBI_TaxID=240292;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=ATCC 29413;
 RG US DOE Joint Genome Institute;
 RA Copeland A., Lucas S., Lapidus A., Barry K., Detter J.C., Glavina T.,
 RA Hammon N., Istrati S., Pittluck S., Saunders E.H., Schmutz J.,
 RA Latimer F., Land M., Kyriades N., Savromatis K., Richardson P.;
 RT "Complete sequence of Anabaena variabilis ATCC 29413";
 CC Submitted (SEP-2005) to the EMBL/GenBank/DBJ databases.
 CC Copyrighted under the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs License
 CC EMBL, CP000117; ABA22265.1; -, Genomic_DNA.
 DR EMBL; CP000117; ABA22265.1; -, Genomic_DNA.
 KW Hypothetical protein; Signal.
 FT SIGNAL 1 29 Potential.
 SQ SEQUENCE 333 AA; 37291 MW; 1F1E00BAF83097F3 CRC64;
 Query Match 84.2%; Score 32; DB 2; Length 333;
 Best Local Similarity 71.4%; Pred. No. 3.1e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 WMLSAFS 7
 Db 17 WMLSAFS 23

RESULT 40
 Q2TRN7_9DIPT PRELIMINARY; PRT; 341 AA.
 ID Q2TRN7_9DIPT
 AC Q2TRN7;
 DT 24-JAN-2006, integrated into UniProtKB/TrEMBL.
 DT 24-JAN-2006, sequence version 1.
 DT 07-MAR-2006, entry version 3.
 DE NADH dehydrogenase subunit 2 (Fragment).
 OS Drosophila unimaculata.
 OC Mitochondrion.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Peerygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=311483;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Wang B.-C., Park J., Watabe H.-A., Gao J.-D., Aotsuka T., Chen H.-W.,

```

RA Zhang Y.-P.;
RT "Phylogenetic analysis of the Drosophila robusta and melanica species
RT groups based on mitochondrial and nuclear DNA sequences.";
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -1- SUBCELLULAR LOCATION: Mitochondrion; mitochondrial inner membrane;
CC multi-pass membrane protein (By similarity).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL: AY750084; AAX52015.1; -; Genomic_DNA.
DR GO: GO:0005739; C:mitochondrion; IEA.
DR Mitochondrion; NAD: Oxidoreductase; Ubiquinone.
FT NON_TER 341 341
SQ SEQUENCE 341 AA; 39719 MW; 4754704028E5995A CRC64;

Query Match 84.2%; Score 32; DB 2; Length 341;
Best Local Similarity 85.7%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 WMLSAFS 7
Db 185 WMLSALS 191

RESULT 41
Q2TNS2_9DIPT PRELIMINARY; PRT; 341 AA.
ID Q2TNS2_9DIPT
AC Q2TNS2_9DIPT
DT 24-JAN-2006, integrated into UniProtKB/TrEMBL.
DT 24-JAN-2006, sequence version 1.
DT 07-MAR-2006, entry version 3.
DE NADH dehydrogenase subunit 2 (Fragment).
OS Drosophila mediocincta.
OC Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=311480;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Wang B.-C., Park J., Watabe H.-A., Gao J.-J., Aotsuka T., Chen H.-W.,
RA Zhang Y.-P.;
RT "Phylogenetic analysis of the Drosophila robusta and melanica species
RT groups based on mitochondrial and nuclear DNA sequences.";
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -1- SUBCELLULAR LOCATION: Mitochondrion; mitochondrial inner membrane;
CC multi-pass membrane protein (By similarity).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL: AY750079; AAX52010.1; -; Genomic_DNA.
DR EMBL: AY750074; AAX52005.1; -; Genomic_DNA.
DR GO: GO:0005739; C:mitochondrion; IEA.
DR Mitochondrion; Ubiquinone.
FT NON_TER 341 341
SQ SEQUENCE 341 AA; 39472 MW; FCA6B7F94E90A733 CRC64;

Query Match 84.2%; Score 32; DB 2; Length 341;
Best Local Similarity 85.7%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 WMLSAFS 7
Db 185 WMLSALS 191

```

```

AC Q2TNS6;
DT 24-JAN-2006, integrated into UniProtKB/TrEMBL.
DT 24-JAN-2006, sequence version 1.
DT 07-MAR-2006, entry version 3.
DE NADH dehydrogenase subunit 2 (Fragment).
OS Drosophila mediocincta.
OC Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=311480;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Wang B.-C., Park J., Watabe H.-A., Gao J.-J., Aotsuka T., Chen H.-W.,
RA Zhang Y.-P.;
RT "Phylogenetic analysis of the Drosophila robusta and melanica species
RT groups based on mitochondrial and nuclear DNA sequences.";
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -1- SUBCELLULAR LOCATION: Mitochondrion; mitochondrial inner membrane;
CC multi-pass membrane protein (By similarity).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL: AY750069; AAX52000.1; -; Genomic_DNA.
DR GO: GO:0005739; C:mitochondrion; IEA.
DR Mitochondrion; NAD: Oxidoreductase; Ubiquinone.
FT NON_TER 341 341
SQ SEQUENCE 341 AA; 39589 MW; A9B3CC084A3B5B95 CRC64;

Query Match 84.2%; Score 32; DB 2; Length 341;
Best Local Similarity 85.7%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 WMLSAFS 7
Db 185 WMLSALS 191

RESULT 43
Q2TNT2_9DIPT PRELIMINARY; PRT; 341 AA.
ID Q2TNT2_9DIPT
AC Q2TNT2_9DIPT
DT 24-JAN-2006, integrated into UniProtKB/TrEMBL.
DT 24-JAN-2006, sequence version 1.
DT 07-MAR-2006, entry version 3.
DE NADH dehydrogenase subunit 2 (Fragment).
OS Drosophila trisgana.
OC Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=65966;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Wang B.-C., Park J., Watabe H.-A., Gao J.-J., Aotsuka T., Chen H.-W.,
RA Zhang Y.-P.;
RT "Phylogenetic analysis of the Drosophila robusta and melanica species
RT groups based on mitochondrial and nuclear DNA sequences.";
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -1- SUBCELLULAR LOCATION: Mitochondrion; mitochondrial inner membrane;
CC multi-pass membrane protein (By similarity).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----

```

Query Match 84.2%; Score 32; DB 2; Length 341;
Best Local Similarity 85.7%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 WMLSAFS 7
|||||
Db 185 WMLSALS 191

RESULT 44

Q2TNT3_9DIPT PRELIMINARY; PRT; 341 AA.
AC Q2TNT3;
DT 24-JAN-2006, integrated into UniProtKB/TrEMBL.
DT 24-JAN-2006, sequence version 1.
DT 07-MAR-2006, entry version 3.
DE NADH dehydrogenase subunit 2 (Fragment).
OS Drosophila paramelanica.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=311477;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Wang B.-C., Park J., Watabe H.-A., Gao J.-J., Aotsuka T., Chen H.-W.,
RA Zhang Y.-P.;
RT "Phylogenetic analysis of the Drosophila robusta and melanica species
groups based on mitochondrial and nuclear DNA sequences.";
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -!- SUBCELLULAR LOCATION: Mitochondrion; mitochondrial inner membrane;
CC multi-pass membrane protein (By similarity).

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs License

CC EMBL: AY750068; AAX5199.1; -!- Genomic_DNA.
DR GO: GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
FT NON_TER 341 341
SQ SEQUENCE 341 AA; 39575 MW; 4A142BDA3BE5BD CRC64;

Query Match 84.2%; Score 32; DB 2; Length 341;
Best Local Similarity 85.7%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 WMLSAFS 7
|||||
Db 185 WMLSALS 191

RESULT 45

Q2TNT5_9DIPT PRELIMINARY; PRT; 341 AA.
AC Q2TNT5;
DT 24-JAN-2006, integrated into UniProtKB/TrEMBL.
DT 24-JAN-2006, sequence version 1.
DT 07-MAR-2006, entry version 3.
DE NADH dehydrogenase subunit 2 (Fragment).
OS Drosophila moritavakii.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=107829;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Wang B.-C., Park J., Watabe H.-A., Gao J.-J., Aotsuka T., Chen H.-W.,
RA Zhang Y.-P.;
RT "Phylogenetic analysis of the Drosophila robusta and melanica species
groups based on mitochondrial and nuclear DNA sequences.";

RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -!- SUBCELLULAR LOCATION: Mitochondrion; mitochondrial inner membrane;
CC multi-pass membrane protein (By similarity).

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs License

CC EMBL: AY750066; AAX51997.1; -!- Genomic_DNA.
DR GO: GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
FT NON_TER 341 341
SQ SEQUENCE 341 AA; 39479 MW; 686D59056FDE76A CRC64;

Query Match 84.2%; Score 32; DB 2; Length 341;
Best Local Similarity 85.7%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 WMLSAFS 7
|||||
Db 185 WMLSALS 191

RESULT 46

Q2TNT6_9DIPT PRELIMINARY; PRT; 341 AA.
AC Q2TNT6;
DT 24-JAN-2006, integrated into UniProtKB/TrEMBL.
DT 24-JAN-2006, sequence version 1.
DT 07-MAR-2006, entry version 3.
DE NADH dehydrogenase subunit 2 (Fragment).
OS Drosophila macromelanica.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=189258;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Wang B.-C., Park J., Watabe H.-A., Gao J.-J., Aotsuka T., Chen H.-W.,
RA Zhang Y.-P.;
RT "Phylogenetic analysis of the Drosophila robusta and melanica species
groups based on mitochondrial and nuclear DNA sequences.";
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -!- SUBCELLULAR LOCATION: Mitochondrion; mitochondrial inner membrane;
CC multi-pass membrane protein (By similarity).

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs License

CC EMBL: AY750065; AAX51996.1; -!- Genomic_DNA.
DR GO: GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
FT NON_TER 341 341
SQ SEQUENCE 341 AA; 39646 MW; 5D1FDBA2786B4A5 CRC64;

Query Match 84.2%; Score 32; DB 2; Length 341;
Best Local Similarity 85.7%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 WMLSAFS 7
|||||
Db 185 WMLSALS 191

RESULT 47

Q2TNT7_DROMX PRELIMINARY; PRT; 341 AA.
AC Q2TNT7;
DT 24-JAN-2006, integrated into UniProtKB/TrEMBL.
DT 24-JAN-2006, sequence version 1.
DT 07-MAR-2006, entry version 3.

DE NADH dehydrogenase subunit 2 (Fragment).
 OS Drosophila melanica (Fruit fly).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7252;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Wang B.-C., Park J., Watabe H.-A., Gao J.-J., Aotsuka T., Chen H.-W.,
 RA Zhang Y.-P.;
 RT "Phylogenetic analysis of the Drosophila robusta and melanica species
 RT groups based on mitochondrial and nuclear DNA sequences.";
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
 CC -1- SUBCELLULAR LOCATION: Mitochondrion; mitochondrial inner membrane;
 CC multi-pass membrane protein (By similarity).
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs License
 CC -----
 CC EMBL: AY75064; AAX51995.1; -; Genomic_DNA.
 DR GO: GO:0005739; C:mitochondrion; IEA.
 KM Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
 FT NON TER 341
 SQ SEQUENCE 341 AA; 39639 MW; BA0B0907D1793705 CRC64;
 Query Match 84.2%; Score 32; DB 2; Length 341;
 Best Local Similarity 85.7%; Pred. No. 3.2e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 WMLSAFS 7
 Db 185 WMLSAFS 191
 RESULT 48
 Q2TNT9_9DIPT PRELIMINARY; PRT; 341 AA.
 ID Q2TNT9;
 AC Q2TNT9;
 DT 24-JAN-2006, integrated into UniProtKB/TrEMBL.
 DT 24-JAN-2006, sequence version 1.
 DT 07-MAR-2006, entry version 3.
 DE NADH dehydrogenase subunit 2 (Fragment).
 OS Drosophila tsigana.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=65966;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Wang B.-C., Park J., Watabe H.-A., Gao J.-J., Aotsuka T., Chen H.-W.,
 RA Zhang Y.-P.;
 RT "Phylogenetic analysis of the Drosophila robusta and melanica species
 RT groups based on mitochondrial and nuclear DNA sequences.";
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
 CC -1- SUBCELLULAR LOCATION: Mitochondrion; mitochondrial inner membrane;
 CC multi-pass membrane protein (By similarity).
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs License
 CC -----
 CC EMBL: AY75062; AAX51993.1; -; Genomic_DNA.
 DR GO: GO:0005739; C:mitochondrion; IEA.
 KM Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
 FT NON TER 341
 SQ SEQUENCE 341 AA; 39512 MW; 65E3CB198B6B59C4 CRC64;
 Query Match 84.2%; Score 32; DB 2; Length 341;
 Best Local Similarity 85.7%; Pred. No. 3.2e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
 Db 185 WMLSAFS 191
 RESULT 49
 Q2TNU1_9DIPT PRELIMINARY; PRT; 341 AA.
 ID Q2TNU1;
 AC Q2TNU1;
 DT 24-JAN-2006, integrated into UniProtKB/TrEMBL.
 DT 24-JAN-2006, sequence version 1.
 DT 07-MAR-2006, entry version 3.
 DE NADH dehydrogenase subunit 2 (Fragment).
 OS Drosophila longiserrata.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=311474;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Wang B.-C., Park J., Watabe H.-A., Gao J.-J., Aotsuka T., Chen H.-W.,
 RA Zhang Y.-P.;
 RT "Phylogenetic analysis of the Drosophila robusta and melanica species
 RT groups based on mitochondrial and nuclear DNA sequences.";
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
 CC -1- SUBCELLULAR LOCATION: Mitochondrion; mitochondrial inner membrane;
 CC multi-pass membrane protein (By similarity).
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs License
 CC -----
 CC EMBL: AY75061; AAX51992.1; -; Genomic_DNA.
 DR GO: GO:0005739; C:mitochondrion; IEA.
 KM Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
 FT NON TER 341
 SQ SEQUENCE 341 AA; 39530 MW; DIE3C9604EB5869 CRC64;
 Query Match 84.2%; Score 32; DB 2; Length 341;
 Best Local Similarity 85.7%; Pred. No. 3.2e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 WMLSAFS 7
 Db 185 WMLSAFS 191
 RESULT 50
 Q2TNU1_9DIPT PRELIMINARY; PRT; 341 AA.
 ID Q2TNU1;
 AC Q2TNU1;
 DT 24-JAN-2006, integrated into UniProtKB/TrEMBL.
 DT 24-JAN-2006, sequence version 1.
 DT 07-MAR-2006, entry version 3.
 DE NADH dehydrogenase subunit 2 (Fragment).
 OS Drosophila tsigana.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=65966;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Wang B.-C., Park J., Watabe H.-A., Gao J.-J., Aotsuka T., Chen H.-W.,
 RA Zhang Y.-P.;
 RT "Phylogenetic analysis of the Drosophila robusta and melanica species
 RT groups based on mitochondrial and nuclear DNA sequences.";
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
 CC -1- SUBCELLULAR LOCATION: Mitochondrion; mitochondrial inner membrane;
 CC multi-pass membrane protein (By similarity).
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs License
 CC -----
 CC EMBL: AY75062; AAX51993.1; -; Genomic_DNA.
 DR GO: GO:0005739; C:mitochondrion; IEA.
 KM Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
 FT NON TER 341
 SQ SEQUENCE 341 AA; 39512 MW; 65E3CB198B6B59C4 CRC64;
 Query Match 84.2%; Score 32; DB 2; Length 341;
 Best Local Similarity 85.7%; Pred. No. 3.2e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL; AY750060; AAX51991.1; -; Genomic_DNA.
DR EMBL; AY750058; AAX51989.1; -; Genomic_DNA.
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
FT NON TER 341 341
SQ SEQUENCE 341 AA; 39530 MW; D1B3C9604E6B5869 CRC64;

Query Match 84.2%; Score 32; DB 2; Length 341;
Best Local Similarity 85.7%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
   |||||
Db 185 WMLSHLS 191
```

Search completed: August 29, 2006, 06:13:05
Job time : 77 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2006, 06:14:29 ; Search time 12.0594 Seconds
(without alignments)
50.808 Million cell updates/sec

Title: US-10-541-343-3
Perfect score: 38
Sequence: 1 WMSAFS 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues
Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 500 summaries

Database : Issued Patents AA: *
1: /EMC_Celerra_SIDS3/prodata/2/iaa/5_COMB.pep.*
2: /EMC_Celerra_SIDS3/prodata/2/iaa/6_COMB.pep.*
3: /EMC_Celerra_SIDS3/prodata/2/iaa/7_COMB.pep.*
4: /EMC_Celerra_SIDS3/prodata/2/iaa/H_COMB.pep.*
5: /EMC_Celerra_SIDS3/prodata/2/iaa/PCTUS_COMB.pep.*
6: /EMC_Celerra_SIDS3/prodata/2/iaa/RE_COMB.pep.*
7: /EMC_Celerra_SIDS3/prodata/2/iaa/backfilest1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	100.0	33	2	US-08-928-662-4
2	38	100.0	33	2	US-09-552-018-4
3	38	100.0	33	2	US-09-989-481-4
4	33	86.8	46	2	US-07-757-022B-18
5	33	86.8	46	3	US-10-124-557-18
6	33	86.8	237	2	US-07-757-022B-72
7	33	86.8	237	2	US-10-124-557-72
8	33	86.8	372	2	US-07-757-022B-64
9	33	86.8	372	2	US-10-124-557-64
10	33	86.8	422	2	US-07-757-022B-68
11	33	86.8	422	2	US-10-124-557-68
12	33	86.8	423	2	US-07-757-022B-66
13	33	86.8	423	2	US-10-124-557-66
14	33	86.8	463	2	US-07-757-022B-54
15	33	86.8	463	2	US-10-124-557-54
16	33	86.8	1270	2	US-07-757-022B-44
17	33	86.8	1270	2	US-10-124-557-44
18	33	86.8	1311	2	US-07-757-022B-42
19	33	86.8	1311	2	US-10-124-557-42
20	33	86.8	1313	2	US-07-757-022B-142
21	33	86.8	1313	2	US-10-124-557-142
22	33	86.8	1314	2	US-07-757-022B-50
23	33	86.8	1314	2	US-10-124-557-50
24	33	86.8	1320	2	US-07-757-022B-46
25	33	86.8	1320	2	US-07-757-022B-60
26	33	86.8	1320	2	US-10-164-595-58

27	33	86.8	1320	3	US-10-124-557-46	Sequence 46, Appl
28	33	86.8	1320	3	US-10-124-557-60	Sequence 60, Appl
29	33	86.8	1354	2	US-07-757-022B-48	Sequence 48, Appl
30	33	86.8	1354	3	US-10-124-557-48	Sequence 48, Appl
31	33	86.8	1361	2	US-07-757-022B-40	Sequence 40, Appl
32	33	86.8	1361	2	US-10-124-557-40	Sequence 40, Appl
33	33	86.8	1363	2	US-07-757-022B-52	Sequence 52, Appl
34	33	86.8	1363	3	US-10-124-557-52	Sequence 52, Appl
35	33	86.8	1404	2	US-07-757-022B-2	Sequence 2, Appl
36	33	86.8	1404	2	US-07-757-022B-62	Sequence 62, Appl
37	33	86.8	1404	2	US-10-164-595-78	Sequence 78, Appl
38	33	86.8	1404	2	US-09-298-970A-1	Sequence 1, Appl
39	33	86.8	1404	2	US-09-897-188-1	Sequence 1, Appl
40	33	86.8	1404	3	US-09-556-246-1	Sequence 1, Appl
41	33	86.8	1404	3	US-10-124-557-2	Sequence 2, Appl
42	33	86.8	1404	3	US-10-124-557-62	Sequence 62, Appl
43	33	86.8	1411	2	US-09-949-016-10827	Sequence 10827, A
44	32	84.2	103	2	US-09-107-532A-4889	Sequence 4889, Ap
45	32	84.2	262	2	US-09-252-991A-22359	Sequence 22359, A
46	32	84.2	280	2	US-09-712-363-283	Sequence 283, Ap
47	32	84.2	437	2	US-09-489-039A-7403	Sequence 7403, Ap
48	31	81.6	229	2	US-09-586-106D-99	Sequence 99, Appl
49	31	81.6	229	2	US-10-799-870-99	Sequence 99, Appl
50	30	78.9	119	2	US-09-107-532A-3700	Sequence 3700, Ap
51	30	78.9	191	2	US-09-107-532A-5382	Sequence 5382, Ap
52	30	78.9	230	2	US-09-252-991A-130553	Sequence 130553, A
53	30	78.9	321	2	US-09-328-352-7627	Sequence 7627, Ap
54	30	78.9	332	2	US-09-080-205-6	Sequence 6, Appl
55	30	78.9	359	2	US-09-489-039A-13566	Sequence 13566, A
56	30	78.9	628	2	US-09-252-991A-24741	Sequence 24741, A
57	29	76.3	127	2	US-08-846-762-20	Sequence 438, Ap
58	29	76.3	127	2	US-08-846-762-20	Sequence 438, Ap
59	29	76.3	303	1	US-08-846-762-91	Sequence 20, Appl
60	29	76.3	303	1	US-08-846-762-91	Sequence 20, Appl
61	29	76.3	335	2	US-10-012-231A-130	Sequence 130, Appl
62	29	76.3	335	2	US-10-015-389A-130	Sequence 130, Appl
63	29	76.3	335	2	US-10-006-768A-130	Sequence 130, Appl
64	29	76.3	335	2	US-10-015-671A-130	Sequence 130, Appl
65	29	76.3	335	2	US-10-015-893A-130	Sequence 130, Appl
66	29	76.3	335	2	US-10-011-833A-130	Sequence 130, Appl
67	29	76.3	335	2	US-10-006-041A-130	Sequence 130, Appl
68	29	76.3	335	2	US-10-012-064A-130	Sequence 130, Appl
69	29	76.3	335	2	US-10-015-392A-130	Sequence 130, Appl
70	29	76.3	335	3	US-10-011-795A-130	Sequence 130, Appl
71	29	76.3	335	3	US-10-015-386A-130	Sequence 130, Appl
72	29	76.3	335	3	US-10-012-121A-130	Sequence 130, Appl
73	29	76.3	335	3	US-10-006-485A-130	Sequence 130, Appl
74	29	76.3	335	3	US-10-006-746A-130	Sequence 130, Appl
75	29	76.3	335	3	US-10-012-752A-130	Sequence 130, Appl
76	29	76.3	335	3	US-10-017-253A-130	Sequence 130, Appl
77	29	76.3	335	3	US-10-015-519A-130	Sequence 130, Appl
78	29	76.3	335	3	US-10-015-715A-130	Sequence 130, Appl
79	29	76.3	335	3	US-10-007-236A-130	Sequence 130, Appl
80	29	76.3	339	2	US-09-252-991A-28730	Sequence 28730, A
81	29	76.3	445	2	US-10-360-101-266	Sequence 266, App
82	29	76.3	771	2	US-09-248-796A-15107	Sequence 15107, A
83	29	76.3	817	2	US-09-248-796A-19402	Sequence 19402, A
84	29	76.3	1037	2	US-09-949-016-6723	Sequence 6723, Ap
85	29	76.3	1061	2	US-09-949-016-10165	Sequence 10165, A
86	29	76.3	1287	1	US-08-477-451-7	Sequence 7, Appl
87	29	76.3	1287	1	US-09-036-987A-5	Sequence 5, Appl
88	29	76.3	4928	2	US-09-370-700-5	Sequence 5, Appl
89	29	76.3	4928	2	US-09-603-207-5	Sequence 5, Appl
90	29	76.3	4928	2	US-10-329-148A-5	Sequence 5, Appl
91	28	73.7	80	2	US-09-673-395A-166	Sequence 366, App
92	28	73.7	89	2	US-10-002-344A-168	Sequence 344, App
93	28	73.7	101	2	US-09-673-395A-601	Sequence 601, App
94	28	73.7	103	2	US-10-094-749-2935	Sequence 2935, App
95	28	73.7	127	2	US-09-311-021-94	Sequence 94, Appl
96	28	73.7	130	2	US-09-513-899C-4265	Sequence 4265, Ap
97	28	73.7	157	2	US-09-107-532A-5501	Sequence 5501, Ap
98	28	73.7	162	2	US-09-711-164-372	Sequence 372, App
99	28	73.7	165	2	US-09-621-976-4112	Sequence 4112, App

100	28	73.7	214	2	US-09-605-703B-2796	Sequence 2796, Ap	173	27	71.1	289	2	US-09-489-039A-10785	Sequence 10785, A
101	28	73.7	223	2	US-09-902-540-16445	Sequence 16445, A	174	27	71.1	294	2	US-09-270-767-44600	Sequence 44800, A
102	28	73.7	237	2	US-09-252-991A-19010	Sequence 19010, A	175	27	71.1	298	2	US-09-248-796A-15967	Sequence 15967, A
103	28	73.7	247	1	US-09-328-352-5648	Sequence 5648, Ap	176	27	71.1	313	2	US-09-362-506-2	Sequence 2, Appl1
104	28	73.7	288	1	US-08-781-562-3	Sequence 3, Appl1	177	27	71.1	339	3	US-09-944-049-6	Sequence 6, Appl1
105	28	73.7	305	2	US-09-866-061-9	Sequence 9, Appl1	178	27	71.1	330	2	US-09-502-540-14418	Sequence 14418, A
106	28	73.7	305	2	US-09-452-937A-44	Sequence 44, Appl1	179	27	71.1	332	2	US-09-270-767-56940	Sequence 56940, A
107	28	73.7	305	2	US-09-664-840-3	Sequence 3, Appl1	180	27	71.1	412	2	US-09-550-772-6	Sequence 6, Appl1
108	28	73.7	306	2	US-09-866-061-10	Sequence 10, Appl1	181	27	71.1	414	2	US-09-328-352-6961	Sequence 6961, Ap
109	28	73.7	322	2	US-09-252-991A-28933	Sequence 28933, A	182	27	71.1	400	2	US-09-248-796A-15811	Sequence 15811, A
110	28	73.7	343	2	US-09-734-237B-35	Sequence 35, Appl1	183	27	71.1	452	2	US-09-348-796A-15424	Sequence 15424, A
111	28	73.7	343	2	US-09-734-237B-37	Sequence 37, Appl1	184	27	71.1	452	2	US-09-602-787A-502	Sequence 502, App
112	28	73.7	381	2	US-09-711-681-2	Sequence 2, Appl1	185	27	71.1	433	2	US-09-252-991A-13015	Sequence 13015, A
113	28	73.7	381	2	US-10-274-266-2	Sequence 2, Appl1	186	27	71.1	548	1	US-09-902-540-12324	Sequence 12324, A
114	28	73.7	388	2	US-09-248-796A-19306	Sequence 19306, A	187	27	71.1	528	1	US-08-403-8520-21	Sequence 21, Appl1
115	28	73.7	402	2	US-09-489-039A-10321	Sequence 10321, A	188	27	71.1	528	2	US-08-510-6466-22	Sequence 22, Appl1
116	28	73.7	407	2	US-09-252-991A-21511	Sequence 21511, A	189	27	71.1	528	2	US-09-231-818-21	Sequence 21, Appl1
117	28	73.7	418	2	US-09-252-991A-29452	Sequence 29452, A	190	27	71.1	528	2	US-09-633-3599-21	Sequence 21, Appl1
118	28	73.7	435	2	US-09-252-991A-18163	Sequence 18163, A	191	27	71.1	578	2	US-09-270-767-41696	Sequence 41696, A
119	28	73.7	452	2	US-09-602-787A-398	Sequence 398, App	192	27	71.1	593	2	US-09-470-767-35149	Sequence 35149, A
120	28	73.7	503	2	US-09-252-991A-17891	Sequence 17891, A	193	27	71.1	593	2	US-09-270-767-50366	Sequence 50366, A
121	28	73.7	510	2	US-10-074-547-2	Sequence 2, Appl1	194	27	71.1	611	1	US-08-677-049-2	Sequence 2, Appl1
122	28	73.7	694	2	US-09-252-991A-16899	Sequence 16899, A	195	27	71.1	647	2	US-08-485-3558-50	Sequence 50, Appl1
123	28	73.7	697	2	US-09-830-433A-52	Sequence 52, Appl1	196	27	71.1	647	2	US-09-194-613-3	Sequence 3, Appl1
124	28	73.7	697	2	US-09-830-433A-77	Sequence 77, Appl1	197	27	71.1	662	3	US-10-114-270-124	Sequence 124, Appl1
125	28	73.7	746	2	US-10-094-749-2678	Sequence 2678, Ap	198	27	71.1	712	3	US-08-485-3558-52	Sequence 52, Appl1
126	28	73.7	908	2	US-09-252-991A-24682	Sequence 24682, A	199	27	71.1	725	2	US-10-199-229-15	Sequence 15, Appl1
127	28	73.7	998	2	US-10-101-464A-931	Sequence 931, App	200	27	71.1	760	2	US-09-489-039A-15061	Sequence 12061, A
128	28	73.7	1085	1	US-08-485-588-5	Sequence 5, Appl1	201	27	71.1	760	2	US-09-543-681A-7842	Sequence 7842, Ap
129	28	73.7	1085	1	US-08-484-585-5	Sequence 5, Appl1	202	27	71.1	760	2	US-09-262-537-10	Sequence 10, Appl1
130	28	73.7	1085	1	US-08-480-751-5	Sequence 5, Appl1	203	27	71.1	879	2	US-09-252-991A-15784	Sequence 19784, A
131	28	73.7	1085	1	US-08-943-986-5	Sequence 5, Appl1	204	27	71.1	890	2	US-09-248-796A-15242	Sequence 19242, A
132	28	73.7	1085	2	US-08-353-784-5	Sequence 5, Appl1	205	27	71.1	996	2	US-10-101-464A-889	Sequence 889, App
133	28	73.7	1085	2	US-08-484-719B-5	Sequence 5, Appl1	206	27	71.1	996	2	US-10-101-464A-933	Sequence 933, App
134	28	73.7	1085	2	US-08-484-159-5	Sequence 5, Appl1	207	27	71.1	1001	1	US-07-797-556-6	Sequence 6, Appl1
135	28	73.7	1089	2	US-10-012-231A-102	Sequence 102, App	208	27	71.1	1001	1	US-07-943-843-2	Sequence 2, Appl1
136	28	73.7	1089	2	US-10-015-389A-102	Sequence 102, App	209	27	71.1	1001	1	US-08-347-003-2	Sequence 2, Appl1
137	28	73.7	1089	2	US-10-006-768A-102	Sequence 102, App	210	27	71.1	1097	1	US-07-943-843-6	Sequence 6, Appl1
138	28	73.7	1089	2	US-10-015-671A-102	Sequence 102, App	211	27	71.1	1097	1	US-08-347-003-6	Sequence 6, Appl1
139	28	73.7	1089	2	US-10-015-393A-102	Sequence 102, App	212	27	71.1	1097	2	US-09-949-016-6209	Sequence 6209, Ap
140	28	73.7	1089	2	US-10-011-833A-102	Sequence 102, App	213	27	71.1	1114	2	US-09-262-537-34	Sequence 34, Appl1
141	28	73.7	1089	2	US-10-006-041A-102	Sequence 102, App	214	27	71.1	1119	2	US-09-949-016-10081	Sequence 10081, A
142	28	73.7	1089	2	US-10-012-064A-102	Sequence 102, App	215	27	71.1	1133	2	US-09-262-537-4	Sequence 4, Appl1
143	28	73.7	1089	2	US-10-015-392A-102	Sequence 102, App	216	27	71.1	1134	2	US-09-538-092-3399	Sequence 399, App
144	28	73.7	1089	3	US-10-011-795B-102	Sequence 102, App	217	27	71.1	1177	2	US-09-262-537-2	Sequence 2, Appl1
145	28	73.7	1089	3	US-10-015-386A-102	Sequence 102, App	218	27	71.1	1207	2	US-09-927-112-2	Sequence 2, Appl1
146	28	73.7	1089	3	US-10-012-121A-102	Sequence 102, App	219	27	71.1	1224	2	US-09-930-872-4	Sequence 4, Appl1
147	28	73.7	1089	3	US-10-006-485A-102	Sequence 102, App	220	27	71.1	1224	2	US-10-217-774-4	Sequence 4, Appl1
148	28	73.7	1089	3	US-10-006-746A-102	Sequence 102, App	221	27	71.1	1403	2	US-09-262-537-6	Sequence 6, Appl1
149	28	73.7	1089	3	US-10-012-752A-102	Sequence 102, App	222	27	71.1	1703	2	US-09-824-574-3	Sequence 340, App
150	28	73.7	1089	3	US-10-017-253A-102	Sequence 102, App	223	27	71.1	1703	2	US-09-487-558B-340	Sequence 340, App
151	28	73.7	1089	3	US-10-015-519A-102	Sequence 102, App	224	26	68.4	14	1	US-08-311-611A-176	Sequence 176, App
152	28	73.7	1089	3	US-10-015-715A-102	Sequence 102, App	225	26	68.4	14	1	US-08-372-783-176	Sequence 176, App
153	28	73.7	1089	3	US-10-007-236A-102	Sequence 102, App	226	26	68.4	14	1	US-08-372-105-176	Sequence 176, App
154	28	73.7	1850	2	US-09-620-093A-5	Sequence 5, Appl1	227	26	68.4	14	1	US-08-306-473A-176	Sequence 176, App
155	28	73.7	2247	2	US-09-091-219-2	Sequence 2, Appl1	228	26	68.4	14	1	US-08-485-445A-176	Sequence 176, App
156	28	73.7	2247	2	US-09-660-541-2	Sequence 2, Appl1	229	26	68.4	14	1	US-08-621-259A-75	Sequence 75, Appl1
157	28	73.7	3340	2	US-09-252-991A-23568	Sequence 23568, A	230	26	68.4	14	2	US-09-119-2693-176	Sequence 176, App
158	27	71.1	69	2	US-09-621-976-5242	Sequence 5242, Ap	231	26	68.4	14	2	US-08-657-162-176	Sequence 176, App
159	27	71.1	69	2	US-09-621-976-5243	Sequence 5243, Ap	232	26	68.4	14	2	US-09-224-480-176	Sequence 176, App
160	27	71.1	70	2	US-09-270-767-60266	Sequence 60266, A	233	26	68.4	14	2	US-09-677-664B-75	Sequence 75, Appl1
161	27	71.1	71	2	US-09-134-000C-6365	Sequence 6365, Ap	234	26	68.4	14	2	US-09-689-097-173	Sequence 173, App
162	27	71.1	72	2	US-09-248-796A-25448	Sequence 25448, A	235	26	68.4	14	5	PCT-US95-00498-176	Sequence 176, App
163	27	71.1	75	2	US-09-621-976-4440	Sequence 4440, Ap	236	26	68.4	14	5	PCT-US95-00656-75	Sequence 75, Appl1
164	27	71.1	115	2	US-10-104-047-3940	Sequence 3940, Ap	237	26	68.4	14	5	PCT-US95-09262-75	Sequence 75, Appl1
165	27	71.1	170	2	US-09-270-767-45761	Sequence 45761, A	238	26	68.4	15	2	US-09-255-501-136	Sequence 136, App
166	27	71.1	176	2	US-09-462-842-3	Sequence 3, Appl1	239	26	68.4	15	2	US-09-255-501-137	Sequence 137, App
167	27	71.1	176	2	US-09-393-171-3	Sequence 3, Appl1	240	26	68.4	15	2	US-09-255-501-138	Sequence 138, App
168	27	71.1	210	2	US-09-543-681A-8168	Sequence 8168, Ap	241	26	68.4	15	2	US-09-060-872A-137	Sequence 137, App
169	27	71.1	210	2	US-09-107-532A-4485	Sequence 4485, Ap	242	26	68.4	15	2	US-09-060-872A-137	Sequence 138, App
170	27	71.1	215	2	US-09-949-016-11620	Sequence 11620, A	243	26	68.4	15	2	US-09-060-872A-138	Sequence 136, App
171	27	71.1	229	2	US-09-252-991A-19517	Sequence 19517, A	244	26	68.4	15	2	US-09-500-135C-136	Sequence 137, App
172	27	71.1	269	2	US-09-902-540-10447	Sequence 10447, A	245	26	68.4	15	2	US-09-500-135C-137	Sequence 137, App

246	26	68.4	15	2	US-09-500-135C-138	Sequence 138, App	319	26	68.4	325	2	US-08-158-735A-13	Sequence 13, Appl
247	26	68.4	15	2	US-09-768-080-136	Sequence 136, App	320	26	68.4	329	2	US-09-270-767-42146	Sequence 42146, A
248	26	68.4	15	2	US-09-768-080-137	Sequence 137, App	321	26	68.4	331	2	US-09-603-208A-242	Sequence 242, App
249	26	68.4	15	2	US-09-768-080-138	Sequence 138, App	322	26	68.4	345	2	US-08-981-700A-5	Sequence 5, Appl
250	26	68.4	15	2	US-09-677-822A-136	Sequence 136, App	323	26	68.4	345	2	US-09-543-681A-1546	Sequence 7546, Ap
251	26	68.4	15	2	US-09-677-822A-137	Sequence 137, App	324	26	68.4	346	2	US-09-199-737-5	Sequence 5, Appl
252	26	68.4	15	2	US-09-677-822A-138	Sequence 138, App	325	26	68.4	346	2	US-08-993-088A-3	Sequence 3, Appl
253	26	68.4	18	2	US-09-407-687-14	Sequence 14, Appl	326	26	68.4	346	2	US-08-993-424B-3	Sequence 3, Appl
254	26	68.4	29	1	US-08-117-952-796	Sequence 796, App	327	26	68.4	346	2	US-09-058-333A-5	Sequence 5, Appl
255	26	68.4	60	2	US-09-621-976-5871	Sequence 5871, Ap	328	26	68.4	346	2	US-09-595-544-6	Sequence 6, Appl
256	26	68.4	61	2	US-09-134-000C-4340	Sequence 4340, Ap	329	26	68.4	346	2	US-08-899-112B-30	Sequence 30, Appl
257	26	68.4	72	2	US-09-248-796A-24979	Sequence 24979, A	330	26	68.4	346	2	US-09-011-553-7	Sequence 7, Appl
258	26	68.4	92	2	US-09-370-767-58901	Sequence 58901, A	331	26	68.4	346	2	US-10-007-132-5	Sequence 5, Appl
259	26	68.4	105	2	US-09-247-155-165	Sequence 165, App	332	26	68.4	347	2	US-09-489-038A-7810	Sequence 7810, Ap
260	26	68.4	105	2	US-09-903-190-165	Sequence 165, App	333	26	68.4	347	2	US-08-513-974B-46	Sequence 46, Appl
261	26	68.4	108	2	US-09-482-873-123	Sequence 123, App	334	26	68.4	348	2	US-08-513-974B-342	Sequence 342, App
262	26	68.4	140	2	US-09-949-016-6875	Sequence 6875, Ap	335	26	68.4	348	2	US-08-993-088A-10	Sequence 10, Appl
263	26	68.4	142	2	US-10-001-254-29	Sequence 29, Appl	336	26	68.4	348	2	US-08-993-424B-10	Sequence 10, Appl
264	26	68.4	147	2	US-09-540-236-3333	Sequence 3333, Ap	337	26	68.4	348	2	US-08-540-650B-2	Sequence 2, Appl
265	26	68.4	147	2	US-09-270-767-43535	Sequence 43535, A	338	26	68.4	348	2	US-09-595-548-5	Sequence 5, Appl
266	26	68.4	147	2	US-09-949-016-7124	Sequence 7124, Ap	339	26	68.4	348	2	US-09-461-438B-46	Sequence 46, Appl
267	26	68.4	162	2	US-09-252-991A-29190	Sequence 29190, A	340	26	68.4	348	2	US-09-603-680-10	Sequence 10, Appl
268	26	68.4	162	2	US-09-621-976-4128	Sequence 4128, Ap	341	26	68.4	349	2	US-08-513-974B-343	Sequence 343, App
269	26	68.4	174	2	US-09-270-767-46937	Sequence 46937, A	342	26	68.4	349	2	US-08-993-088A-11	Sequence 11, Appl
270	26	68.4	175	2	US-09-267-963D-32	Sequence 32, Appl	343	26	68.4	349	2	US-08-993-424B-11	Sequence 11, Appl
271	26	68.4	176	2	US-09-267-963D-31	Sequence 31, Appl	344	26	68.4	349	2	US-08-540-650B-5	Sequence 5, Appl
272	26	68.4	177	2	US-09-270-767-57414	Sequence 57414, A	345	26	68.4	349	2	US-08-693-308-2	Sequence 2, Appl
273	26	68.4	179	2	US-09-252-991A-24524	Sequence 24524, A	346	26	68.4	349	2	US-09-595-549-7	Sequence 7, Appl
274	26	68.4	181	2	US-09-902-540-9914	Sequence 9914, Ap	347	26	68.4	349	2	US-09-603-680-11	Sequence 11, Appl
275	26	68.4	182	2	US-09-267-963D-30	Sequence 30, Appl	348	26	68.4	349	2	US-08-981-700A-6	Sequence 6, Appl
276	26	68.4	184	2	US-09-252-991A-19139	Sequence 19139, A	349	26	68.4	349	2	US-09-826-509-503	Sequence 503, App
277	26	68.4	188	2	US-09-328-352-4862	Sequence 4862, Ap	350	26	68.4	351	2	US-08-513-974B-344	Sequence 344, App
278	26	68.4	196	1	US-08-467-822-42	Sequence 42, Appl	351	26	68.4	351	1	US-08-153-848-24	Sequence 24, Appl
279	26	68.4	196	2	US-09-107-383-11	Sequence 11, Appl	352	26	68.4	352	1	US-09-299-843A-24	Sequence 24, Appl
280	26	68.4	196	2	US-08-466-248-42	Sequence 42, Appl	353	26	68.4	353	2	US-09-336-643A-2	Sequence 2, Appl
281	26	68.4	196	2	US-09-643-914-11	Sequence 11, Appl	354	26	68.4	353	2	US-09-336-643A-2	Sequence 2, Appl
282	26	68.4	196	2	US-09-742-361A-11	Sequence 11, Appl	355	26	68.4	355	5	PCR-US93-11153-24	Sequence 24, Appl
283	26	68.4	197	2	US-09-489-039A-14194	Sequence 14194, A	356	26	68.4	356	1	US-08-946-875-8	Sequence 8, Appl
284	26	68.4	197	2	US-09-270-767-42732	Sequence 42732, A	357	26	68.4	357	1	US-08-102-385G-8	Sequence 8, Appl
285	26	68.4	197	2	US-09-248-796A-26798	Sequence 26798, A	358	26	68.4	358	2	US-09-298-337B-66	Sequence 66, Appl
286	26	68.4	199	1	US-08-467-822-43	Sequence 43, Appl	359	26	68.4	359	2	US-09-488-265B-18	Sequence 18, Appl
287	26	68.4	199	2	US-08-432-697-43	Sequence 43, Appl	360	26	68.4	360	2	US-09-949-016-9465	Sequence 9465, Ap
288	26	68.4	199	2	US-08-466-248-43	Sequence 43, Appl	361	26	68.4	361	2	US-08-900-230-5	Sequence 5, Appl
289	26	68.4	211	2	US-09-902-540-10687	Sequence 10687, A	362	26	68.4	362	2	US-09-605-703B-1020	Sequence 1020, Ap
290	26	68.4	219	2	US-09-388-221B-21	Sequence 21, Appl	363	26	68.4	363	2	US-09-605-703B-1022	Sequence 1022, Ap
291	26	68.4	230	2	US-09-064-703-7	Sequence 7, Appl	364	26	68.4	364	2	US-09-605-703B-1022	Sequence 1022, Ap
292	26	68.4	235	2	US-09-345-236B-108	Sequence 108, App	365	26	68.4	365	2	US-09-684-855-119	Sequence 119, App
293	26	68.4	235	2	US-09-248-796A-15600	Sequence 15600, A	366	26	68.4	366	2	US-09-488-265B-18	Sequence 18, Appl
294	26	68.4	241	2	US-09-469-039A-7888	Sequence 7888, Ap	367	26	68.4	367	2	US-09-252-991A-17309	Sequence 17309, A
295	26	68.4	246	2	US-09-270-767-44455	Sequence 44455, A	368	26	68.4	368	2	US-08-660-347-2	Sequence 2, Appl
296	26	68.4	249	2	US-09-489-039A-9413	Sequence 9413, Ap	369	26	68.4	369	2	US-08-993-359-28	Sequence 28, Appl
297	26	68.4	255	2	US-09-248-796A-19627	Sequence 19627, A	370	26	68.4	370	2	US-09-482-558A-28	Sequence 28, Appl
298	26	68.4	253	2	US-09-107-532A-7230	Sequence 7230, Ap	371	26	68.4	371	2	US-09-273-871A-4	Sequence 4, Appl
299	26	68.4	255	2	US-10-094-749-2763	Sequence 2763, Ap	372	26	68.4	372	2	US-09-482-558A-28	Sequence 28, Appl
300	26	68.4	255	2	US-09-270-767-37246	Sequence 37246, A	373	26	68.4	373	2	US-09-538-092-1002	Sequence 1002, Ap
301	26	68.4	276	2	US-09-270-767-52463	Sequence 52463, A	374	26	68.4	374	2	US-09-603-208A-240	Sequence 240, App
302	26	68.4	279	2	US-09-134-000C-5300	Sequence 5300, Ap	375	26	68.4	375	2	US-09-328-352-4776	Sequence 4726, Ap
303	26	68.4	280	2	US-09-255-501-209	Sequence 209, App	376	26	68.4	376	2	US-09-215-418-2	Sequence 2, Appl
304	26	68.4	280	2	US-09-060-872A-209	Sequence 209, App	377	26	68.4	377	2	US-09-453-681A-5855	Sequence 5855, Ap
305	26	68.4	280	2	US-09-500-135C-209	Sequence 209, App	378	26	68.4	378	2	US-09-107-532A-4860	Sequence 4860, Ap
306	26	68.4	280	2	US-09-768-080-209	Sequence 209, App	379	26	68.4	379	2	US-09-134-000C-65335	Sequence 6535, Ap
307	26	68.4	280	2	US-09-677-822A-209	Sequence 209, App	380	26	68.4	380	2	US-09-270-767-57120	Sequence 57120, A
308	26	68.4	281	2	US-09-973-278-172	Sequence 172, App	381	26	68.4	381	2	US-09-886-954A-1	Sequence 1, Appl
309	26	68.4	281	2	US-09-973-278-172	Sequence 172, App	382	26	68.4	382	2	US-09-291-023A-2	Sequence 2, Appl
310	26	68.4	281	2	US-09-973-278-172	Sequence 172, App	383	26	68.4	383	2	US-09-291-023A-4	Sequence 4, Appl
311	26	68.4	293	2	US-09-252-991A-30174	Sequence 30174, A	384	26	68.4	384	2	US-09-291-023A-14	Sequence 14, Appl
312	26	68.4	308	2	US-10-014-269-31	Sequence 31, Appl	385	26	68.4	385	2	US-09-291-023A-15	Sequence 15, Appl
313	26	68.4	308	2	US-10-002-974-31	Sequence 31, Appl	386	26	68.4	386	2	US-09-540-715A-2	Sequence 2, Appl
314	26	68.4	314	2	US-08-981-857D-13	Sequence 13, Appl	387	26	68.4	387	2	US-09-540-715A-4	Sequence 4, Appl
315	26	68.4	314	2	US-09-982-704-13	Sequence 13, Appl	388	26	68.4	388	2	US-09-540-715A-14	Sequence 14, Appl
316	26	68.4	318	2	US-09-355-166-16	Sequence 16, Appl	389	26	68.4	389	2	US-09-540-715A-15	Sequence 15, Appl
317	26	68.4	320	2	US-09-248-796A-20977	Sequence 20977, A	390	26	68.4	390	2	US-09-540-715A-15	Sequence 15, Appl
318	26	68.4	323	2	US-08-158-735A-12	Sequence 12, Appl	391	26	68.4	391	2	US-09-949-016-7855	Sequence 7855, Ap

APPLICANT: Chau, Raymond M.W.
TITLE OF INVENTION: Isolation and Use of Motoneuronotropic Factors
FILE REFERENCE: 12592-3
CURRENT APPLICATION NUMBER: US/09/592,018
CURRENT FILING DATE: 2000-06-12
PRIOR APPLICATION NUMBER: US 08/9928862
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: US 08/751225
PRIOR FILING DATE: 1996-11-15
PRIOR APPLICATION NUMBER: US 60/026792
PRIOR FILING DATE: 1996-09-27
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn version 3.0
SEQ ID NO 4
LENGTH: 33
TYPE: PRT
ORGANISM: Homo Sapiens
US-09-592-018-4

Query Match 100.0%; Score 38; DB 2; Length 33;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
DB 12 WMLSAFS 18

RESULT 3
US-09-989-481-4
Sequence 4; Application US/09989481
Patent No. 6841531
GENERAL INFORMATION:
APPLICANT: Chau, Raymond M.W.
TITLE OF INVENTION: Isolation and Use of Motoneuronotropic Factors
FILE REFERENCE: 12592-4
CURRENT APPLICATION NUMBER: US/09/989,481
CURRENT FILING DATE: 2001-11-20
PRIOR APPLICATION NUMBER: US 09/633,447
PRIOR FILING DATE: 2000-08-07
PRIOR APPLICATION NUMBER: US 08/9928862
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: US 08/751225
PRIOR FILING DATE: 1996-11-15
PRIOR APPLICATION NUMBER: US 60/026792
PRIOR FILING DATE: 1996-09-27
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn version 3.0
SEQ ID NO 4
LENGTH: 33
TYPE: PRT
ORGANISM: Homo Sapiens
US-09-989-481-4

Query Match 100.0%; Score 38; DB 2; Length 33;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
DB 12 WMLSAFS 18

RESULT 4
US-07-757-022B-18
Sequence 18; Application US/07757022B
Patent No. 6431142
GENERAL INFORMATION:
APPLICANT: Gesner, Thomas G.
APPLICANT: Clark, Stephen C.
APPLICANT: Turner, Katherine
APPLICANT: Hewick, Rodney M.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors

NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/757,022B
FILING DATE: 19910910
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Gesner, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)876-1170
TELEFAX: (617)876-5851
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-757-022B-18

Query Match 86.8%; Score 33; DB 2; Length 46;
Best Local Similarity 85.7%; Pred. No. 15;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
DB 5 WMLSAFS 11

RESULT 5
US-10-124-557-18
Sequence 18; Application US/10124557
Patent No. 7030223
GENERAL INFORMATION:
APPLICANT: Turner, Katherine
APPLICANT: Clark, Stephen C.
APPLICANT: Jacobs, Kenneth
APPLICANT: Hewick, Rodney M.
APPLICANT: Gesner, Thomas G.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/124,557
FILING DATE: 16-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Cseert, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)876-1170
TELEFAX: (617)876-5851
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-10-124-557-18

Query Match 86.8%; Score 33; DB 3; Length 46;
Best Local Similarity 85.7%; Pred. No. 15;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
Db 5 WMLSPFS 11

RESULT 6
US-07-757-022B-72
Sequence 72, Application US/07757022B
Patent No. 6433142
GENERAL INFORMATION:
APPLICANT: Gesner, Thomas G.
APPLICANT: Clark, Stephen C.
APPLICANT: Turner, Katherine
APPLICANT: Hewick, Rodney M.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/757,022B
FILING DATE: 19910910
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Cseert, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)876-1170
TELEFAX: (617)876-5851
INFORMATION FOR SEQ ID NO: 72:
SEQUENCE CHARACTERISTICS:
LENGTH: 237 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-757-022B-72

Query Match 86.8%; Score 33; DB 2; Length 237;
Best Local Similarity 85.7%; Pred. No. 77;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
Db 111 WMLSPFS 117

RESULT 7
US-10-124-557-72
Sequence 72, Application US/10124557
Patent No. 7030223
GENERAL INFORMATION:
APPLICANT: Turner, Katherine
APPLICANT: Clark, Stephen C.
APPLICANT: Jacobs, Kenneth
APPLICANT: Hewick, Rodney M.
APPLICANT: Gesner, Thomas G.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/124,557
FILING DATE: 16-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Cseert, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190

TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)876-1170
TELEFAX: (617)876-5851
INFORMATION FOR SEQ ID NO: 72:
SEQUENCE CHARACTERISTICS:
LENGTH: 237 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 72:
US-10-124-557-72

Query Match 86.8%; Score 33; DB 3; Length 237;
Best Local Similarity 85.7%; Pred. No. 77;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSPFS 7
Db 111 WMLSPFS 117

RESULT 8
US-07-757-022B-64
Sequence 64, Application US/07757022B
Patent No. 6431142
GENERAL INFORMATION:
APPLICANT: Gesner, Thomas G.
APPLICANT: Clark, Stephen C.
APPLICANT: Turner, Katherine
APPLICANT: Hewick, Rodney M.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/757,022B
FILING DATE: 19910910
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Cserr, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)876-1170
TELEFAX: (617)876-5851
INFORMATION FOR SEQ ID NO: 64:
SEQUENCE CHARACTERISTICS:
LENGTH: 372 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein

US-07-757-022B-64

Query Match 86.8%; Score 33; DB 2; Length 372;
Best Local Similarity 85.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSPFS 7
Db 139 WMLSPFS 145

RESULT 9
US-10-124-557-64
Sequence 64, Application US/10124557
Patent No. 7030223
GENERAL INFORMATION:
APPLICANT: Turner, Katherine
APPLICANT: Clark, Stephen C.
APPLICANT: Jacobs, Kenneth
APPLICANT: Hewick, Rodney M.
APPLICANT: Gesner, Thomas G.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/124,557
FILING DATE: 16-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Cserr, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)876-1170
TELEFAX: (617)876-5851
INFORMATION FOR SEQ ID NO: 64:
SEQUENCE CHARACTERISTICS:
LENGTH: 372 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 64:
US-10-124-557-64

Query Match 86.8%; Score 33; DB 3; Length 372;
Best Local Similarity 85.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSPFS 7
Db 139 WMLSPFS 145

```
RESULT 10
US-07-757-022B-68
; Sequence 68, Application US/07757022B
; Patent No. 6433142
; GENERAL INFORMATION:
; APPLICANT: Gesner, Thomas G.
; APPLICANT: Clark, Stephen C.
; APPLICANT: Turner, Katherine
; APPLICANT: Hewick, Rodney M.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/757,022B
; FILING DATE: 19910910
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cserr, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)876-1170
; TELEFAX: (617)876-5851
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 422 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-757-022B-68

Query Match      86.8%; Score 33; DB 2; Length 422;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```

; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cserr, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)876-1170
; TELEFAX: (617)876-5851
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 422 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 68:
;
; US-10-124-557-68

Query Match      86.8%; Score 33; DB 3; Length 422;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Oy      1 WMLSAFS 7
Db      189 WMLSPFS 195

RESULT 11
US-10-124-557-68
; Sequence 68, Application US/10124557
; Patent No. 7030223
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; APPLICANT: Clark, Stephen C.
; APPLICANT: Jacobs, Kenneth
; APPLICANT: Hewick, Rodney M.
; APPLICANT: Gesner, Thomas G.
```

```

; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cserr, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)876-1170
; TELEFAX: (617)876-5851
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 422 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 68:
;
; US-10-124-557-68

Query Match      86.8%; Score 33; DB 3; Length 422;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/757,022B
FILING DATE: 19910910
CLASSIFICATION: 530
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Cserr, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)876-1170
TELEFAX: (617)876-5851
INFORMATION FOR SEQ ID NO: 66:
SEQUENCE CHARACTERISTICS:
LENGTH: 423 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-757-022B-66

Query Match 86.8%; Score 33; DB 2; Length 423;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSPFS 7
DB 190 WMLSPFS 196

RESULT 13
US-10-124-557-66
Sequence 66; Application US/10124557
Patent No. 7030223
GENERAL INFORMATION:
APPLICANT: Turner, Katherine
Clark, Stephen C.
Jacobs, Kenneth
Hewick, Rodney M.
Gesner, Thomas G.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/124,557
FILING DATE: 16-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991

APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Cserr, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)876-1170
TELEFAX: (617)876-5851
INFORMATION FOR SEQ ID NO: 66:
SEQUENCE CHARACTERISTICS:
LENGTH: 423 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 66:
US-10-124-557-66

Query Match 86.8%; Score 33; DB 3; Length 423;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSPFS 7
DB 190 WMLSPFS 196

RESULT 14
US-07-757-022B-54
Sequence 54; Application US/07757022B
Patent No. 6431142
GENERAL INFORMATION:
APPLICANT: Gesner, Thomas G.
Clark, Stephen C.
Turner, Katherine
Hewick, Rodney M.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/757,022B
FILING DATE: 19910910
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Cserr, Luann
REGISTRATION NUMBER: 31,822

REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 876-1170
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 463 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-757-022B-54

Query Match 86.8%; Score 33; DB 2; Length 463;
Best Local Similarity 85.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
Db 230 WMLSPFS 236

RESULT 15
US-10-124-557-54
Sequence 54, Application US/10124557
Patent No. 7030223
GENERAL INFORMATION:
APPLICANT: Turner, Katherine
Clark, Stephen C.
Jacobs, Kenneth
Hewick, Rodney M.
Gesner, Thomas G.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/124,557
FILING DATE: 16-Apr-2002
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Geert, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 876-1170
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 463 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 54:
US-10-124-557-54

Query Match 86.8%; Score 33; DB 3; Length 463;
Best Local Similarity 85.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
Db 230 WMLSPFS 236

RESULT 16
US-07-757-022B-44
Sequence 44, Application US/07757022B
Patent No. 643142
GENERAL INFORMATION:
APPLICANT: Gesner, Thomas G.
Clark, Stephen C.
APPLICANT: Turner, Katherine
Hewick, Rodney M.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/757,022B
FILING DATE: 19910910
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Geert, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 876-1170
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 1270 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-757-022B-44

Query Match 86.8%; Score 33; DB 2; Length 1270;
Best Local Similarity 85.7%; Pred. No. 4.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
Db 1037 WMLSPFS 1043

RESULT 17
US-10-124-557-44
; Sequence 44, Application US/10124557
; Patent No. 7030223
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Gesner, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cseert, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 876-1170
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1270 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 44:
US-10-124-557-44
Query Match 86.8%; Score 33; DB 3; Length 1270;
Best Local Similarity 85.7%; Pred. No. 4.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
Db 1037 WMLSPFS 1043

RESULT 18
US-07-757-022B-42
; Sequence 42, Application US/07757022B
; Patent No. 6433142
; GENERAL INFORMATION:
; APPLICANT: Gesner, Thomas G.
; Clark, Stephen C.
; APPLICANT: Turner, Katherine
; Hewick, Rodney M.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143

CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/757,022B
; FILING DATE: 19910910
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cseert, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 876-1170
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1311 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-757-022B-42
Query Match 86.8%; Score 33; DB 2; Length 1311;
Best Local Similarity 85.7%; Pred. No. 4.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
Db 1078 WMLSPFS 1084

RESULT 19
US-10-124-557-42
; Sequence 42, Application US/10124557
; Patent No. 7030223
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Gesner, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/124,557
FILING DATE: 16-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Cserr, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)876-1170
TELEFAX: (617)876-5851
INFORMATION FOR SEQ ID NO: 42:
SEQUENCE CHARACTERISTICS:
LENGTH: 1311 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 42:
US-10-124-557-42

Query Match 86.8%; Score 33; DB 3; Length 1311;
Best Local Similarity 85.7%; Pred. No. 4.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
Db 1078 WMLSPFS 1084

RESULT 20
US-07-757-022B-142
Sequence 142, Application US/07/57022B
Patent No. 6433142
GENERAL INFORMATION:
APPLICANT: Gesner, Thomas G.
APPLICANT: Clark, Stephen C.
APPLICANT: Turner, Katherine
APPLICANT: Hewick, Rodney M.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/757,022B
FILING DATE: 19910910
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/546,114

FILING DATE: 29-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Cserr, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)876-1170
TELEFAX: (617)876-5851
INFORMATION FOR SEQ ID NO: 142:
SEQUENCE CHARACTERISTICS:
LENGTH: 1313 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-757-022B-142

Query Match 86.8%; Score 33; DB 2; Length 1313;
Best Local Similarity 85.7%; Pred. No. 4.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
Db 1080 WMLSPFS 1086

RESULT 21
US-10-124-557-142
Sequence 142, Application US/10/124557
Patent No. 7030223
GENERAL INFORMATION:
APPLICANT: Turner, Katherine
APPLICANT: Clark, Stephen C.
APPLICANT: Jacobs, Kenneth
APPLICANT: Hewick, Rodney M.
APPLICANT: Gesner, Thomas G.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/124,557
FILING DATE: 16-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Cserr, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 876-1170
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 142:
SEQUENCE CHARACTERISTICS:
LENGTH: 1313 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 142:
US-10-124-557-142

Query Match 86.8%; Score 33; DB 3; Length 1313;
Best Local Similarity 85.7%; Pred. No. 4.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSPFS 7
DB 1080 WMLSPFS 1086

RESULT 22
US-07-757-022B-50
Sequence 50, Application US/07757022B
Patent No. 6433142
GENERAL INFORMATION:
APPLICANT: Geener, Thomas G.
APPLICANT: Clark, Stephen C.
APPLICANT: Turner, Katherine
APPLICANT: Hewick, Rodney M.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/757,022B
FILING DATE: 19910910
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Geert, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 876-1170
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 1314 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-757-022B-50

Query Match 86.8%; Score 33; DB 2; Length 1314;
Best Local Similarity 85.7%; Pred. No. 4.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSPFS 7
DB 1081 WMLSPFS 1087

RESULT 23
US-10-124-557-50
Sequence 50, Application US/10124557
Patent No. 7030223
GENERAL INFORMATION:
APPLICANT: Turner, Katherine
APPLICANT: Clark, Stephen C.
APPLICANT: Jacobs, Kenneth
APPLICANT: Hewick, Rodney M.
APPLICANT: Geener, Thomas G.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/124,557
FILING DATE: 16-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Geert, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 876-1170
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 1314 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 50:
US-10-124-557-50

Query Match 86.8%; Score 33; DB 3; Length 1314;
Best Local Similarity 85.7%; Pred. No. 4.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSPFS 7
DB 1081 WMLSPFS 1087

RESULT 24

US-07-757-022B-46
; Sequence 46, Application US/07757022B
; Patent No. 6433142
; GENERAL INFORMATION:
; APPLICANT: Gesner, Thomas G.
; APPLICANT: Clark, Stephen C.
; APPLICANT: Turner, Katherine
; APPLICANT: Hewick, Rodney M.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/757,022B
; FILING DATE: 19910910
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cserr, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)876-1170
; TELEFAX: (617)876-5851
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1320 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-757-022B-46

Query Match 86.8%; Score 33; DB 2; Length 1320;
Best Local Similarity 85.7%; Pred. No. 4.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 WMLSPFS 7
Db 1087 WMLSPFS 1093

RESULT 25
US-07-757-022B-60
; Sequence 60, Application US/07757022B
; Patent No. 6433142
; GENERAL INFORMATION:
; APPLICANT: Gesner, Thomas G.
; APPLICANT: Clark, Stephen C.
; APPLICANT: Turner, Katherine
; APPLICANT: Hewick, Rodney M.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143

CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/757,022B
; FILING DATE: 19910910
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cserr, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)876-1170
; TELEFAX: (617)876-5851
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1320 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-757-022B-60

Query Match 86.8%; Score 33; DB 2; Length 1320;
Best Local Similarity 85.7%; Pred. No. 4.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 WMLSPFS 7
Db 1087 WMLSPFS 1093

RESULT 26
US-10-164-595-58
; Sequence 58, Application US/10164595
; Patent No. 6657054
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies, Inc
; TITLE OF INVENTION: Regulated Angiogenesis Genes and Polypeptides
; FILE REFERENCE: IU 103 R1
; CURRENT APPLICATION NUMBER: US/10/164,595
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 58
; LENGTH: 1320
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-164-595-58

Query Match 86.8%; Score 33; DB 2; Length 1320;
Best Local Similarity 85.7%; Pred. No. 4.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
Db 1087 WMLSPFS 1093

RESULT 27

US-10-124-557-46
; Sequence 46, Application US/10124557
; Patent No. 7030223
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Gesner, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; PRIORITY:
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cseerr, Luanu
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)876-1170
; TELEFAX: (617)876-5851
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1320 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 46:
US-10-124-557-46

Query Match 86.8%; Score 33; DB 3; Length 1320;
Best Local Similarity 85.7%; Pred. No. 4.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
Db 1087 WMLSPFS 1093

RESULT 28
US-10-124-557-60
; Sequence 60, Application US/10124557
; Patent No. 7030223
; GENERAL INFORMATION:

APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Gesner, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; PRIORITY:
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cseerr, Luanu
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)876-1170
; TELEFAX: (617)876-5851
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1320 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 60:
US-10-124-557-60

Query Match 86.8%; Score 33; DB 3; Length 1320;
Best Local Similarity 85.7%; Pred. No. 4.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
Db 1087 WMLSPFS 1093

RESULT 29
US-07-757-022B-48
; Sequence 48, Application US/07757022B
; Patent No. 6433142
; GENERAL INFORMATION:
; APPLICANT: Gesner, Thomas G.
; APPLICANT: Clark, Stephen C.
; APPLICANT: Turner, Katherine
; APPLICANT: Hewick, Rodney M.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts

COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/757,022B
FILING DATE: 19910910
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Cserr, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 876-1170
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 1354 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-757-022B-48

Query Match 86.8%; Score 33; DB 2; Length 1354;
Best Local Similarity 85.7%; Pred. No. 4.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSPFS 7
Db 1121 WMLSPFS 1127

RESULT 30
US-10-124-557-48
Sequence 48, Application US/10124557
Patent No. 7030223
GENERAL INFORMATION:
APPLICANT: Turner, Katherine
Clark, Stephen C.
Jacobs, Kenneth
Hewick, Rodney M.
Gesner, Thomas G.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/124,557

FILING DATE: 16-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Cserr, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 876-1170
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 1354 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 48:
US-10-124-557-48

Query Match 86.8%; Score 33; DB 3; Length 1354;
Best Local Similarity 85.7%; Pred. No. 4.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSPFS 7
Db 1121 WMLSPFS 1127

RESULT 31
US-07-757-022B-40
Sequence 40, Application US/07757022B
Patent No. 6433142
GENERAL INFORMATION:
APPLICANT: Gesner, Thomas G.
Clark, Stephen C.
APPLICANT: Turner, Katherine
Hewick, Rodney M.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/757,022B
FILING DATE: 19910910
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Geert, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)876-1170
TELEFAX: (617)876-5851
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 1361 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-757-022B-40

Query Match 86.8%; Score 33; DB 2; Length 1361;
Best Local Similarity 85.7%; Pred. No. 4.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 WMLSAFS 7
Db 1128 WMLSPFS 1134

RESULT 32
US-10-124-557-40
Sequence 40, Application US/10124557
Patent No. 7030223
GENERAL INFORMATION:
APPLICANT: Turner, Katherine
Clark, Stephen C.
Jacobs, Kenneth
Hewick, Rodney M.
Geener, Thomas G.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 CambridgePark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/124,557
FILING DATE: 16-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Geert, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)876-1170
TELEFAX: (617)876-5851
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 1361 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 40:
US-10-124-557-40

Query Match 86.8%; Score 33; DB 3; Length 1361;
Best Local Similarity 85.7%; Pred. No. 4.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 WMLSAFS 7
Db 1128 WMLSPFS 1134

RESULT 33
US-07-757-022B-52
Sequence 52, Application US/07757022B
Patent No. 6433142
GENERAL INFORMATION:
APPLICANT: Geener, Thomas G.
APPLICANT: Clark, Stephen C.
APPLICANT: Turner, Katherine
APPLICANT: Hewick, Rodney M.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 CambridgePark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/757,022B
FILING DATE: 19910910
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Geert, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)876-1170
TELEFAX: (617)876-5851
INFORMATION FOR SEQ ID NO: 52:
SEQUENCE CHARACTERISTICS:
LENGTH: 1363 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-757-022B-52

Query Match 86.8%; Score 33; DB 2; Length 1363;
Best Local Similarity 85.7%; Pred. No. 4.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
Db 1130 WMLSPFS 1136

RESULT 34

US-10-124-557-52
; Sequence 52, Application US/10124557
; Patent No. 7030223
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Gesner, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cserr, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)876-1170
; TELEFAX: (617)876-5851
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1363 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 52:
US-10-124-557-52

Query Match 86.8%; Score 33; DB 3; Length 1363;
Best Local Similarity 85.7%; Pred. No. 4.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
Db 1130 WMLSPFS 1136

RESULT 35

US-07-757-022B-2
; Sequence 2, Application US/07757022B
; Patent No. 6433142
; GENERAL INFORMATION:
; APPLICANT: Gesner, Thomas G.

APPLICANT: Clark, Stephen C.
APPLICANT: Turner, Katherine
APPLICANT: Hewick, Rodney M.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/757,022B
FILING DATE: 19910910
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Cserr, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)876-1170
TELEFAX: (617)876-5851
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1404 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-757-022B-2

Query Match 86.8%; Score 33; DB 2; Length 1404;
Best Local Similarity 85.7%; Pred. No. 4.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
Db 1171 WMLSPFS 1177

RESULT 36

US-07-757-022B-62
; Sequence 62, Application US/07757022B
; Patent No. 6433142
; GENERAL INFORMATION:
; APPLICANT: Gesner, Thomas G.
; APPLICANT: Clark, Stephen C.
; APPLICANT: Turner, Katherine
; APPLICANT: Hewick, Rodney M.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts


```

; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/757,022B
; FILING DATE: 19910910
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cseert, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)876-1170
; TELEFAX: (617)876-5851
; INFORMATION FOR SEQ ID NO: 62:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1404 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-757-022B-62

Query Match      86.8%; Score 33; DB 2; Length 1404;
Best Local Similarity 85.7%; Pred. No. 4.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 WMLSAFS 7
Db      1171 WMLSPFS 1177

RESULT 37
US-10-164-595-78
; Sequence 78, Application US/10164595
; Patent No. 6657054
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies, Inc
; TITLE OF INVENTION: Regulated Angiogenesis Genes and Polypeptides
; FILE REFERENCE: 1U 103 R1
; CURRENT APPLICATION NUMBER: US/10/164,595
; CURRENT FILING DATE: 2002-06-10
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 78
; LENGTH: 1404
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-164-595-78

Query Match      86.8%; Score 33; DB 2; Length 1404;
Best Local Similarity 85.7%; Pred. No. 4.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 WMLSAFS 7
Db      1171 WMLSPFS 1177
```

```

; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/757,022B
; FILING DATE: 19910910
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cseert, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)876-1170
; TELEFAX: (617)876-5851
; INFORMATION FOR SEQ ID NO: 62:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1404 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-757-022B-62

Query Match      86.8%; Score 33; DB 2; Length 1404;
Best Local Similarity 85.7%; Pred. No. 4.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 WMLSAFS 7
Db      1171 WMLSPFS 1177

RESULT 38
US-09-298-970A-1
; Sequence 1, Application US/09298970A
; Patent No. 6743774
; GENERAL INFORMATION:
; APPLICANT: Jay, Gregory D.
; TITLE OF INVENTION: TRIBONECTINS
; FILE REFERENCE: 21486-026
; CURRENT APPLICATION NUMBER: US/09/298,970A
; CURRENT FILING DATE: 2001-06-19
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1404
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-298-970A-1

Query Match      86.8%; Score 33; DB 2; Length 1404;
Best Local Similarity 85.7%; Pred. No. 4.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 WMLSAFS 7
Db      1171 WMLSPFS 1177

RESULT 39
US-09-897-188-1
; Sequence 1, Application US/09897188
; Patent No. 6960562
; GENERAL INFORMATION:
; APPLICANT: Jay, Gregory D.
; TITLE OF INVENTION: Tribonectin Polypeptides and Uses Thereof
; FILE REFERENCE: 21486-026 CIP2
; CURRENT APPLICATION NUMBER: US/09/897,188
; CURRENT FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: 09/298,970
; PRIOR FILING DATE: 1999-04-23
; PRIOR APPLICATION NUMBER: 09/556,246
; PRIOR FILING DATE: 2000-04-24
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1404
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-897-188-1

Query Match      86.8%; Score 33; DB 2; Length 1404;
Best Local Similarity 85.7%; Pred. No. 4.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 WMLSAFS 7
Db      1171 WMLSPFS 1177

RESULT 40
US-09-556-246-1
; Sequence 1, Application US/09556246
; Patent No. 7001881
; GENERAL INFORMATION:
; APPLICANT: Gregory D. Jay
; TITLE OF INVENTION: tribonectins
; FILE REFERENCE: 21486-026CIP
; CURRENT APPLICATION NUMBER: US/09/556,246
; CURRENT FILING DATE: 2000-04-24
; PRIOR APPLICATION NUMBER: USSN 09/298/970
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Patentin Ver. 2.0
```

SEQ ID NO 1
LENGTH: 1404
TYPE: PRT
ORGANISM: Homo sapiens
US-09-556-246-1

Query Match 86.8%; Score 33; DB 3; Length 1404;
Best Local Similarity 85.7%; Pred. No. 4.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
Db 1171 WMLSPFS 1177

RESULT 41

US-10-124-557-2
Sequence 2, Application US/10124557
Patent No. 7030223

GENERAL INFORMATION:

APPLICANT: Turner, Katherine
Clark, Stephen C.
Jacobs, Kenneth
Hewick, Rodney M.
Gesner, Thomas G.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/124,557
FILING DATE: 16-Apr-2002
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

REGISTRATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Cserr, Luann

REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)876-1170
TELEFAX: (617)876-5851

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 1404 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-124-557-2

Query Match 86.8%; Score 33; DB 3; Length 1404;
Best Local Similarity 85.7%; Pred. No. 4.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7

Db 1171 WMLSPFS 1177

RESULT 42

US-10-124-557-62
Sequence 62, Application US/10124557
Patent No. 7030223

GENERAL INFORMATION:

APPLICANT: Turner, Katherine
Clark, Stephen C.
Jacobs, Kenneth
Hewick, Rodney M.
Gesner, Thomas G.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/124,557
FILING DATE: 16-Apr-2002
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

REGISTRATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Cserr, Luann

REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)876-1170
TELEFAX: (617)876-5851

INFORMATION FOR SEQ ID NO: 62:

SEQUENCE CHARACTERISTICS:
LENGTH: 1404 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 62:
US-10-124-557-62

Query Match 86.8%; Score 33; DB 3; Length 1404;
Best Local Similarity 85.7%; Pred. No. 4.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
Db 1171 WMLSPFS 1177

RESULT 43

US-09-949-016-10827
Sequence 10827, Application US/09949016
Patent No. 6812339

GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10827
LENGTH: 1411
TYPE: PRT
ORGANISM: Human
US-09-949-016-10827

Query Match
Best Local Similarity 86.8%; Score 33; DB 2; Length 1411;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSPS 7
Db 1178 WMLSPS 1184

RESULT 44
US-09-107-532A-4889
Sequence 4889, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 4889:
SEQUENCE CHARACTERISTICS:
LENGTH: 103 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:

NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...103
SEQUENCE DESCRIPTION: SEQ ID NO: 4889:
US-09-107-532A-4889

Query Match
Best Local Similarity 84.2%; Score 32; DB 2; Length 103;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WMLSP 6
Db 17 WMLSP 22

RESULT 45
US-09-252-991A-22359
Sequence 22359, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 22359
LENGTH: 262
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
FEATURE:
NAME/KEY: UNSURE
LOCATION: (182)
OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
US-09-252-991A-22359

Query Match
Best Local Similarity 84.2%; Score 32; DB 2; Length 262;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSPS 7
Db 234 WMLSPS 240

RESULT 46
US-09-712-363-283
Sequence 283, Application US/09712363
Patent No. 6892139
GENERAL INFORMATION:
APPLICANT: Eisenberg, David
APPLICANT: Rotstein, Sergio H.
APPLICANT: Marcotte, Edward M.
TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND
INTERACTIONS OF PROTEINS BY COMPARATIVE ANALYSIS
FILE REFERENCE: 07419-032001
CURRENT APPLICATION NUMBER: US/09/712,363
CURRENT FILING DATE: 2000-11-13
PRIOR APPLICATION NUMBER: PCT/US00/02246
PRIOR FILING DATE: 2000-01-28
PRIOR APPLICATION NUMBER: 60/179,531
PRIOR FILING DATE: 2000-02-01
PRIOR APPLICATION NUMBER: 60/117,844
PRIOR FILING DATE: 1999-01-29
PRIOR APPLICATION NUMBER: 60/118,206
PRIOR FILING DATE: 1999-02-01
PRIOR APPLICATION NUMBER: 60/126,593
PRIOR FILING DATE: 1999-03-26
PRIOR APPLICATION NUMBER: 60/134,093

```
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/134,092
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/165,124
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/165,086
; PRIOR FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 292
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 283
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-712-363-283

Query Match
Best Local Similarity 84.2%; Score 32; DB 2; Length 280;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
Db 262 WMLAFA 268

RESULT 47
US-09-489-039A-7403
; Sequence 7403, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709,2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 66/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 7403
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7403

Query Match
Best Local Similarity 84.2%; Score 32; DB 2; Length 437;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
Db 391 WMLSAFS 397

RESULT 48
US-09-586-106D-99
; Sequence 99, Application US/09586106D
; Patent No. 6720479
; GENERAL INFORMATION:
; APPLICANT: Wright, David A.
; APPLICANT: Voytas, Daniel F.
; TITLE OF INVENTION: PLANT RETROELEMENTS AND METHODS RELATED THERETO
; FILE REFERENCE: P-1065A
; CURRENT APPLICATION NUMBER: US/09/586,106D
; CURRENT FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/087,125
; PRIOR FILING DATE: 1998-05-29
; PRIOR APPLICATION NUMBER: 09/322,478
; PRIOR FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 99
; LENGTH: 229
; TYPE: PRT
```

```
; ORGANISM: Lycopersicon esculentum
US-09-586-106D-99

Query Match
Best Local Similarity 81.6%; Score 31; DB 2; Length 229;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
Db 144 WMLSAFS 150

RESULT 49
US-10-799-870-99
; Sequence 99, Application US/10799870
; Patent No. 6949695
; GENERAL INFORMATION:
; APPLICANT: Wright, David A.
; APPLICANT: Voytas, Daniel F.
; TITLE OF INVENTION: PLANT RETROELEMENTS AND METHODS RELATED THERETO
; FILE REFERENCE: P-1065A
; CURRENT APPLICATION NUMBER: US/10/799,870
; PRIOR FILING DATE: 2004-03-12
; PRIOR APPLICATION NUMBER: US/09/586,106
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/087,125
; PRIOR FILING DATE: 1998-05-29
; PRIOR APPLICATION NUMBER: 09/322,478
; PRIOR FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 99
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Lycopersicon esculentum
US-10-799-870-99

Query Match
Best Local Similarity 81.6%; Score 31; DB 2; Length 229;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
Db 144 WMLSAFS 150

RESULT 50
US-09-107-532A-3700
; Sequence 3700, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCIT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
```

APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781) 893-5007
TELEFAX: (781) 893-8277
INFORMATION FOR SEQ ID NO: 3700:
SEQUENCE CHARACTERISTICS:
LENGTH: 119 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1..119
SEQUENCE DESCRIPTION: SEQ ID NO: 3700:
US-09-107-532A-3700

Query Match 78.9%; Score 30; DB 2; Length 119;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 WMLSAFS 7
Db 11 WMTSFS 17

Search completed: August 29, 2006, 06:17:48
Job time : 17.0594 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2006, 09:58:08 ; Search time 84.6931 Seconds
(without alignments)
38.285 Million cell updates/sec

Title: US-10-541-343-3

Perfect score: 38

Sequence: 1 WMLSARS 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database :

Published Applications AA Main:*

- 1: /EMC_Celerra_SIDS3/prodata/2/pubppaa/US07_PUBCOMB.pep:*
- 2: /EMC_Celerra_SIDS3/prodata/2/pubppaa/US08_PUBCOMB.pep:*
- 3: /EMC_Celerra_SIDS3/prodata/2/pubppaa/US09_PUBCOMB.pep:*
- 4: /EMC_Celerra_SIDS3/prodata/2/pubppaa/US10_PUBCOMB.pep:*
- 5: /EMC_Celerra_SIDS3/prodata/2/pubppaa/US10B_PUBCOMB.pep:*
- 6: /EMC_Celerra_SIDS3/prodata/2/pubppaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	100.0	7	5	US-10-541-343-3
2	38	100.0	13	5	US-10-541-343-6
3	38	100.0	33	3	US-09-989-481-4
4	38	100.0	33	5	US-10-858-545-4
5	38	100.0	33	5	US-10-858-286-4
6	38	100.0	33	5	US-10-858-543-4
7	38	100.0	33	5	US-10-858-144-4
8	38	100.0	33	5	US-10-541-343-1
9	35	92.1	1760	6	US-11-097-143-16893
10	34	89.5	103	4	US-10-424-599-187845
11	33	86.8	46	4	US-10-124-557-18
12	33	86.8	46	6	US-11-169-232-18
13	33	86.8	61	3	US-09-925-101-1452
14	33	86.8	237	6	US-10-124-557-72
15	33	86.8	237	6	US-11-169-232-72
16	33	86.8	372	4	US-10-124-557-64
17	33	86.8	372	6	US-11-169-232-64
18	33	86.8	422	4	US-10-124-557-68
19	33	86.8	422	6	US-11-169-232-68
20	33	86.8	423	6	US-10-124-557-66
21	33	86.8	423	6	US-11-169-232-66
22	33	86.8	463	4	US-10-124-557-54
23	33	86.8	463	6	US-11-169-232-54
24	33	86.8	538	4	US-10-038-694-4
25	33	86.8	792	3	US-09-802-207-27
26	33	86.8	1046	4	US-10-425-115-37015
27	33	86.8	1270	4	US-10-124-557-44

28	33	86.8	1270	6	US-11-169-232-44	Sequence 44, Appl
29	33	86.8	1311	4	US-10-124-557-42	Sequence 42, Appl
30	33	86.8	1311	6	US-11-169-232-42	Sequence 42, Appl
31	33	86.8	1313	4	US-10-124-557-142	Sequence 142, Appl
32	33	86.8	1313	6	US-11-169-232-142	Sequence 142, Appl
33	33	86.8	1314	4	US-10-124-557-50	Sequence 50, Appl
34	33	86.8	1314	6	US-11-169-232-50	Sequence 50, Appl
35	33	86.8	1320	4	US-10-124-557-46	Sequence 46, Appl
36	33	86.8	1320	6	US-11-169-232-46	Sequence 46, Appl
37	33	86.8	1320	5	US-10-124-557-60	Sequence 60, Appl
38	33	86.8	1320	6	US-11-169-232-58	Sequence 58, Appl
39	33	86.8	1320	6	US-11-169-232-46	Sequence 46, Appl
40	33	86.8	1354	4	US-10-124-557-48	Sequence 48, Appl
41	33	86.8	1354	6	US-11-169-232-48	Sequence 48, Appl
42	33	86.8	1361	4	US-10-124-557-40	Sequence 40, Appl
43	33	86.8	1361	6	US-11-169-232-40	Sequence 40, Appl
44	33	86.8	1363	4	US-10-124-557-52	Sequence 52, Appl
45	33	86.8	1363	6	US-11-169-232-52	Sequence 52, Appl
46	33	86.8	1404	3	US-09-802-207-30	Sequence 30, Appl
47	33	86.8	1404	3	US-09-897-188-1	Sequence 1, Appl
48	33	86.8	1404	4	US-10-124-557-2	Sequence 2, Appl
49	33	86.8	1404	4	US-10-124-557-62	Sequence 62, Appl
50	33	86.8	1404	5	US-10-858-595-1	Sequence 1, Appl
51	33	86.8	1404	5	US-10-868-577A-55	Sequence 55, Appl
52	33	86.8	1404	5	US-10-717-665-78	Sequence 78, Appl
53	33	86.8	1404	6	US-11-169-232-2	Sequence 2, Appl
54	33	86.8	1404	6	US-11-169-232-62	Sequence 62, Appl
55	32	84.2	211	3	US-09-791-932-97	Sequence 97, Appl
56	32	84.2	211	5	US-10-980-388-97	Sequence 97, Appl
57	32	84.2	280	3	US-09-712-363-283	Sequence 283, Appl
58	32	84.2	289	4	US-10-407-960-2	Sequence 2, Appl
59	32	84.2	323	3	US-09-791-932-119	Sequence 119, Appl
60	32	84.2	323	4	US-10-332-156-119	Sequence 1, Appl
61	32	84.2	323	5	US-10-980-388-119	Sequence 119, Appl
62	32	84.2	371	4	US-10-471-115-1	Sequence 1, Appl
63	32	84.2	465	5	US-10-626-832-25	Sequence 25, Appl
64	32	84.2	511	4	US-10-437-963-120104	Sequence 120104, Appl
65	32	84.2	561	5	US-10-505-486-107	Sequence 107, Appl
66	32	84.2	623	4	US-10-437-963-144350	Sequence 144350, Appl
67	32	84.2	747	4	US-10-437-963-144356	Sequence 144356, Appl
68	32	84.2	1100	4	US-10-437-963-178971	Sequence 178971, Appl
69	32	84.2	1923	5	US-10-450-763-51701	Sequence 51701, Appl
70	31	81.6	76	3	US-09-764-891-3265	Sequence 3265, Appl
71	31	81.6	76	5	US-10-501-282-3808	Sequence 3808, Appl
72	31	81.6	77	4	US-10-424-599-184690	Sequence 184690, Appl
73	31	81.6	85	4	US-10-424-599-249435	Sequence 249435, Appl
74	31	81.6	91	4	US-10-001-857-125	Sequence 125, Appl
75	31	81.6	93	4	US-10-425-115-309566	Sequence 309566, Appl
76	31	81.6	99	4	US-10-425-115-260762	Sequence 260762, Appl
77	31	81.6	104	4	US-10-424-599-242038	Sequence 242038, Appl
78	31	81.6	133	4	US-10-425-115-193472	Sequence 193472, Appl
79	31	81.6	146	3	US-09-925-297-643	Sequence 643, Appl
80	31	81.6	148	3	US-09-916-790-30	Sequence 30, Appl
81	31	81.6	148	4	US-10-678-786-30	Sequence 30, Appl
82	31	81.6	163	4	US-10-767-701-59679	Sequence 59679, Appl
83	31	81.6	165	4	US-10-425-115-331668	Sequence 331668, Appl
84	31	81.6	178	4	US-10-156-761-7648	Sequence 7648, Appl
85	31	81.6	182	4	US-10-425-115-213595	Sequence 213595, Appl
86	31	81.6	186	4	US-10-425-115-21364	Sequence 21364, Appl
87	31	81.6	211	4	US-10-425-115-217396	Sequence 217396, Appl
88	31	81.6	229	4	US-10-395-607-99	Sequence 99, Appl
89	31	81.6	229	4	US-10-799-870-99	Sequence 99, Appl
90	31	81.6	263	3	US-09-888-615-96	Sequence 96, Appl
91	31	81.6	263	6	US-11-037-443-96	Sequence 96, Appl
92	31	81.6	302	4	US-10-425-115-213613	Sequence 213613, Appl
93	31	81.6	305	4	US-10-425-115-43471	Sequence 43471, Appl
94	31	81.6	322	4	US-10-424-599-281141	Sequence 281141, Appl
95	31	81.6	352	6	US-11-097-143-38151	Sequence 38151, Appl
96	31	81.6	371	3	US-09-981-900B-14	Sequence 14, Appl
97	31	81.6	373	4	US-10-425-115-211851	Sequence 211851, Appl
98	31	81.6	633	4	US-10-437-963-111838	Sequence 111838, Appl
99	30	78.9	41	5	US-10-948-571-77	Sequence 77, Appl
100	30	78.9	50	4	US-10-029-386-29222	Sequence 29222, Appl

101	30	78.9	51	6	US-11-096-568A-27324	Sequence 27324, A	174	29	76.3	127	3	US-09-933-767-438	Sequence 438, App
102	30	78.9	60	4	US-10-425-115-357009	Sequence 357009, A	175	29	76.3	127	4	US-10-004-860-438	Sequence 438, App
103	30	78.9	71	4	US-10-437-963-023604	Sequence 203604, A	176	29	76.3	137	4	US-10-023-282-438	Sequence 438, App
104	30	78.9	74	4	US-10-424-599-176425	Sequence 176425, A	177	29	76.3	128	4	US-10-425-115-246890	Sequence 246890, A
105	30	78.9	83	4	US-10-767-701-39437	Sequence 39437, A	178	29	76.3	131	4	US-10-437-963-132576	Sequence 132576, A
106	30	78.9	85	4	US-10-437-963-132576	Sequence 132576, A	179	29	76.3	131	4	US-10-437-963-132576	Sequence 132576, A
107	30	78.9	126	4	US-10-425-115-559538	Sequence 259538, A	180	29	76.3	144	4	US-10-767-701-46974	Sequence 46974, A
108	30	78.9	152	3	US-09-974-879-608	Sequence 408, App	181	29	76.3	144	6	US-11-079-463-9208	Sequence 9208, App
109	30	78.9	152	3	US-09-305-736-408	Sequence 408, App	182	29	76.3	177	4	US-10-767-701-46504	Sequence 46504, A
110	30	78.9	152	3	US-09-818-683-408	Sequence 408, App	183	29	76.3	185	5	US-10-719-993-556	Sequence 556, App
111	30	78.9	152	3	US-09-818-683-408	Sequence 408, App	184	29	76.3	184	4	US-10-425-114-53585	Sequence 53585, A
112	30	78.9	152	3	US-10-621-401-192	Sequence 408, App	185	29	76.3	214	4	US-10-425-115-300451	Sequence 300451, A
113	30	78.9	155	5	US-10-501-282-3886	Sequence 3886, App	186	29	76.3	232	4	US-10-437-963-158163	Sequence 158163, A
114	30	78.9	174	3	US-09-974-879-192	Sequence 192, App	187	29	76.3	233	4	US-10-335-977-5070	Sequence 5070, App
115	30	78.9	174	3	US-09-305-736-192	Sequence 192, App	188	29	76.3	230	5	US-10-873-467-36	Sequence 36, App
116	30	78.9	174	3	US-09-818-683-192	Sequence 192, App	189	29	76.3	232	4	US-10-425-114-62075	Sequence 62075, A
117	30	78.9	174	3	US-09-818-683-192	Sequence 192, App	190	29	76.3	232	4	US-10-425-114-62075	Sequence 62075, A
118	30	78.9	174	3	US-10-621-401-192	Sequence 192, App	191	29	76.3	231	6	US-11-188-298-4579	Sequence 4579, App
119	30	78.9	179	4	US-10-767-701-14862	Sequence 54862, A	192	29	76.3	222	2	US-08-834-705-14	Sequence 14, App
120	30	78.9	182	4	US-10-369-493-5876	Sequence 5876, App	193	29	76.3	229	2	US-10-335-977-5071	Sequence 5071, App
121	30	78.9	267	5	US-10-732-923-16550	Sequence 16550, A	194	29	76.3	233	5	US-10-450-763-37146	Sequence 37146, A
122	30	78.9	267	5	US-10-732-923-16551	Sequence 16551, A	195	29	76.3	303	4	US-10-216-209-20	Sequence 20, App
123	30	78.9	267	6	US-11-188-288-20218	Sequence 20218, A	196	29	76.3	303	4	US-10-216-209-91	Sequence 91, App
124	30	78.9	268	6	US-11-188-288-1175	Sequence 1175, App	197	29	76.3	311	6	US-11-079-463-8727	Sequence 8727, App
125	30	78.9	273	4	US-10-369-493-3391	Sequence 3391, App	198	29	76.3	315	4	US-10-424-599-179218	Sequence 179218, A
126	30	78.9	320	6	US-11-112-882-76	Sequence 4, App	199	29	76.3	335	3	US-09-946-374-130	Sequence 4, App
127	30	78.9	320	6	US-11-112-882-76	Sequence 4, App	200	29	76.3	335	3	US-09-374-046A-4	Sequence 4, App
128	30	78.9	326	6	US-11-079-463-6243	Sequence 6243, App	201	29	76.3	335	4	US-10-327-884-170	Sequence 170, App
129	30	78.9	332	4	US-10-260-877-34	Sequence 24, App	202	29	76.3	335	4	US-10-320-163-130	Sequence 130, App
130	30	78.9	332	4	US-10-267-854-6	Sequence 6, App	203	29	76.3	335	4	US-10-006-856A-130	Sequence 130, App
131	30	78.9	433	6	US-11-096-666-10520	Sequence 10520, A	204	29	76.3	335	4	US-10-230-338-170	Sequence 170, App
132	30	78.9	498	4	US-10-156-761-11728	Sequence 11728, A	205	29	76.3	335	4	US-10-218-631-170	Sequence 170, App
133	30	78.9	504	4	US-10-437-963-302303	Sequence 202303, A	206	29	76.3	335	4	US-10-230-414-170	Sequence 170, App
134	30	78.9	593	5	US-10-732-923-8486	Sequence 8486, App	207	29	76.3	335	4	US-10-006-818A-130	Sequence 130, App
135	30	78.9	736	4	US-10-425-114-39293	Sequence 39293, A	208	29	76.3	335	4	US-10-006-485A-130	Sequence 130, App
136	30	78.9	761	6	US-11-188-298-18066	Sequence 18066, A	209	29	76.3	335	4	US-10-013-907A-130	Sequence 130, App
137	30	78.9	818	4	US-10-425-114-41341	Sequence 41341, A	210	29	76.3	335	4	US-10-015-499A-130	Sequence 130, App
138	30	78.9	832	4	US-10-424-599-254916	Sequence 254916, A	211	29	76.3	335	4	US-10-232-224-170	Sequence 170, App
139	30	78.9	832	5	US-10-732-923-8518	Sequence 8518, App	212	29	76.3	335	4	US-10-015-393A-130	Sequence 130, App
140	30	78.9	837	4	US-10-425-115-345033	Sequence 345033, A	213	29	76.3	335	4	US-10-216-159A-170	Sequence 170, App
141	30	78.9	844	5	US-10-732-923-8376	Sequence 8376, App	214	29	76.3	335	4	US-10-015-869A-130	Sequence 130, App
142	30	78.9	847	5	US-10-732-923-8521	Sequence 8521, App	215	29	76.3	335	4	US-10-012-121A-130	Sequence 130, App
143	30	78.9	907	6	US-11-013-314-42	Sequence 42, App	216	29	76.3	335	4	US-10-218-849-170	Sequence 170, App
144	30	78.9	907	6	US-11-097-143-15312	Sequence 15312, A	217	29	76.3	335	4	US-10-227-873-170	Sequence 170, App
145	30	78.9	1087	4	US-10-437-963-129691	Sequence 129691, A	218	29	76.3	335	4	US-10-227-883-170	Sequence 170, App
146	30	78.9	1321	4	US-10-369-493-6076	Sequence 6076, App	219	29	76.3	335	4	US-10-219-076-170	Sequence 170, App
147	30	78.9	1372	4	US-10-369-493-6077	Sequence 6077, App	220	29	76.3	335	4	US-10-230-434-170	Sequence 170, App
148	30	78.9	1769	4	US-10-361-522-3	Sequence 3, App	221	29	76.3	335	4	US-10-006-116A-130	Sequence 130, App
149	30	78.9	2250	5	US-10-922-232B-57	Sequence 57, App	222	29	76.3	335	4	US-10-006-117A-130	Sequence 130, App
150	30	78.9	4025	5	US-10-437-963-193926	Sequence 193926, A	223	29	76.3	335	4	US-10-017-527A-130	Sequence 130, App
151	30	78.9	4060	5	US-10-922-232B-55	Sequence 55, App	224	29	76.3	335	4	US-10-013-913A-130	Sequence 130, App
152	30	78.9	6738	5	US-10-922-232B-56	Sequence 56, App	225	29	76.3	335	4	US-10-219-003-170	Sequence 170, App
153	29	76.3	53	4	US-10-425-115-309164	Sequence 309164, A	226	29	76.3	335	4	US-10-219-075-170	Sequence 170, App
154	29	76.3	62	4	US-10-424-599-168965	Sequence 168965, A	227	29	76.3	335	4	US-10-219-64A-170	Sequence 170, App
155	29	76.3	64	4	US-10-424-599-161491	Sequence 161491, A	228	29	76.3	335	4	US-10-219-466-170	Sequence 170, App
156	29	76.3	75	4	US-10-425-115-15484	Sequence 15484, A	229	29	76.3	335	4	US-10-219-479-170	Sequence 170, App
157	29	76.3	77	4	US-10-425-115-193653	Sequence 193653, A	230	29	76.3	335	4	US-10-219-481-170	Sequence 170, App
158	29	76.3	77	4	US-10-335-977-8349	Sequence 8349, App	231	29	76.3	335	4	US-10-230-260-170	Sequence 170, App
159	29	76.3	79	4	US-10-767-701-46821	Sequence 46821, A	232	29	76.3	335	4	US-10-232-231-170	Sequence 170, App
160	29	76.3	80	4	US-10-767-701-53279	Sequence 53279, A	233	29	76.3	335	4	US-10-232-233-170	Sequence 170, App
161	29	76.3	87	3	US-09-764-860-347	Sequence 347, App	234	29	76.3	335	4	US-10-007-194A-130	Sequence 130, App
162	29	76.3	87	4	US-10-074-095-347	Sequence 347, App	235	29	76.3	335	4	US-10-013-430A-130	Sequence 130, App
163	29	76.3	87	4	US-10-212-872-347	Sequence 347, App	236	29	76.3	335	4	US-10-216-165-170	Sequence 170, App
164	29	76.3	91	4	US-10-424-599-328363	Sequence 328363, A	237	29	76.3	335	4	US-10-218-956-170	Sequence 170, App
165	29	76.3	92	4	US-10-425-115-246889	Sequence 246889, A	238	29	76.3	335	4	US-10-219-468-170	Sequence 170, App
166	29	76.3	94	4	US-10-424-599-203952	Sequence 203952, A	239	29	76.3	335	4	US-10-219-478-170	Sequence 170, App
167	29	76.3	106	4	US-10-264-237-1766	Sequence 1766, App	240	29	76.3	335	4	US-10-219-536-170	Sequence 170, App
168	29	76.3	113	3	US-09-948-783-161	Sequence 161, App	241	29	76.3	335	4	US-10-233-205-170	Sequence 170, App
169	29	76.3	113	3	US-10-425-114-62076	Sequence 62076, A	242	29	76.3	335	4	US-10-011-671A-130	Sequence 130, App
170	29	76.3	114	3	US-09-892-877-160	Sequence 160, App	243	29	76.3	335	4	US-10-012-755A-130	Sequence 130, App
171	29	76.3	117	4	US-10-425-114-61328	Sequence 61328, A	244	29	76.3	335	4	US-10-219-072-170	Sequence 170, App
172	29	76.3	120	4	US-10-425-115-345091	Sequence 345091, A	245	29	76.3	335	4	US-10-219-470-170	Sequence 170, App
173	29	76.3	125	4	US-10-425-115-344975	Sequence 344975, A	246	29	76.3	335	4	US-10-219-474-170	Sequence 170, App

247	29	76.3	335	4	US-10-219-524-170	Sequence 170, App	320	29	76.3	335	4	US-10-017-306A-130	Sequence 130, App
248	29	76.3	335	4	US-10-219-528-170	Sequence 170, App	321	29	76.3	335	4	US-10-017-867A-130	Sequence 130, App
249	29	76.3	335	4	US-10-227-880-170	Sequence 170, App	322	29	76.3	335	4	US-10-012-064A-130	Sequence 130, App
250	29	76.3	335	4	US-10-227-881-170	Sequence 170, App	323	29	76.3	335	4	US-10-013-909A-130	Sequence 130, App
251	29	76.3	335	4	US-10-227-882-170	Sequence 170, App	324	29	76.3	335	4	US-10-015-677A-130	Sequence 130, App
252	29	76.3	335	4	US-10-230-436-170	Sequence 170, App	325	29	76.3	335	4	US-10-015-610A-130	Sequence 130, App
253	29	76.3	335	4	US-10-232-223-170	Sequence 170, App	326	29	76.3	335	4	US-10-223-081-190	Sequence 190, App
254	29	76.3	335	4	US-10-232-225-170	Sequence 170, App	327	29	76.3	335	4	US-10-012-137A-130	Sequence 130, App
255	29	76.3	335	4	US-10-232-227-170	Sequence 170, App	328	29	76.3	335	4	US-10-012-752A-130	Sequence 130, App
256	29	76.3	335	4	US-10-232-229-170	Sequence 170, App	329	29	76.3	335	4	US-10-012-754A-130	Sequence 130, App
257	29	76.3	335	4	US-10-232-234-170	Sequence 170, App	330	29	76.3	335	4	US-10-013-910A-130	Sequence 130, App
258	29	76.3	335	4	US-10-015-386A-130	Sequence 130, App	331	29	76.3	335	4	US-10-013-911A-130	Sequence 130, App
259	29	76.3	335	4	US-10-219-060-170	Sequence 170, App	332	29	76.3	335	4	US-10-013-912A-130	Sequence 130, App
260	29	76.3	335	4	US-10-223-085-190	Sequence 190, App	333	29	76.3	335	4	US-10-015-653A-130	Sequence 130, App
261	29	76.3	335	4	US-10-216-160-170	Sequence 170, App	334	29	76.3	335	4	US-10-218-765-170	Sequence 170, App
262	29	76.3	335	4	US-10-216-162-170	Sequence 170, App	335	29	76.3	335	4	US-10-219-063-170	Sequence 170, App
263	29	76.3	335	4	US-10-216-164-170	Sequence 170, App	336	29	76.3	335	4	US-10-219-066-170	Sequence 170, App
264	29	76.3	335	4	US-10-216-167-170	Sequence 170, App	337	29	76.3	335	4	US-10-219-068-170	Sequence 170, App
265	29	76.3	335	4	US-10-216-168-170	Sequence 170, App	338	29	76.3	335	4	US-10-219-069-170	Sequence 170, App
266	29	76.3	335	4	US-10-219-065-170	Sequence 170, App	339	29	76.3	335	4	US-10-219-073-170	Sequence 170, App
267	29	76.3	335	4	US-10-219-071-170	Sequence 170, App	340	29	76.3	335	4	US-10-219-475-170	Sequence 170, App
268	29	76.3	335	4	US-10-219-074-170	Sequence 170, App	341	29	76.3	335	4	US-10-219-480-170	Sequence 170, App
269	29	76.3	335	4	US-10-219-077-170	Sequence 170, App	342	29	76.3	335	4	US-10-219-532-170	Sequence 170, App
270	29	76.3	335	4	US-10-219-465-170	Sequence 170, App	343	29	76.3	335	4	US-10-219-483-170	Sequence 170, App
271	29	76.3	335	4	US-10-219-467-170	Sequence 170, App	344	29	76.3	335	4	US-10-219-525-170	Sequence 170, App
272	29	76.3	335	4	US-10-219-469-170	Sequence 170, App	345	29	76.3	335	4	US-10-219-526-170	Sequence 170, App
273	29	76.3	335	4	US-10-219-471-170	Sequence 170, App	346	29	76.3	335	4	US-10-219-530-170	Sequence 170, App
274	29	76.3	335	4	US-10-219-473-170	Sequence 170, App	347	29	76.3	335	4	US-10-219-531-170	Sequence 170, App
275	29	76.3	335	4	US-10-219-476-170	Sequence 170, App	348	29	76.3	335	4	US-10-219-532-170	Sequence 170, App
276	29	76.3	335	4	US-10-219-482-170	Sequence 170, App	349	29	76.3	335	4	US-10-219-533-170	Sequence 170, App
277	29	76.3	335	4	US-10-227-874-170	Sequence 170, App	350	29	76.3	335	4	US-10-230-437-170	Sequence 170, App
278	29	76.3	335	4	US-10-227-876-170	Sequence 170, App	351	29	76.3	335	4	US-10-232-228-170	Sequence 170, App
279	29	76.3	335	4	US-10-227-878-170	Sequence 170, App	352	29	76.3	335	4	US-10-012-101B-130	Sequence 130, App
280	29	76.3	335	4	US-10-229-974-170	Sequence 170, App	353	29	76.3	335	4	US-10-015-480A-130	Sequence 130, App
281	29	76.3	335	4	US-10-230-024-170	Sequence 170, App	354	29	76.3	335	4	US-10-015-478A-130	Sequence 130, App
282	29	76.3	335	4	US-10-230-113-170	Sequence 170, App	355	29	76.3	335	4	US-10-223-082-190	Sequence 190, App
283	29	76.3	335	4	US-10-230-183-170	Sequence 170, App	356	29	76.3	335	4	US-10-013-906A-130	Sequence 130, App
284	29	76.3	335	4	US-10-230-234-170	Sequence 170, App	357	29	76.3	335	4	US-10-015-388A-130	Sequence 130, App
285	29	76.3	335	4	US-10-230-306-170	Sequence 170, App	358	29	76.3	335	4	US-10-012-753A-130	Sequence 130, App
286	29	76.3	335	4	US-10-230-426-170	Sequence 170, App	359	29	76.3	335	4	US-10-015-385A-130	Sequence 130, App
287	29	76.3	335	4	US-10-230-427-170	Sequence 170, App	360	29	76.3	335	4	US-10-015-385A-130	Sequence 130, App
288	29	76.3	335	4	US-10-230-433-170	Sequence 170, App	361	29	76.3	335	4	US-10-007-236A-130	Sequence 130, App
289	29	76.3	335	4	US-10-230-435-170	Sequence 170, App	362	29	76.3	335	4	US-10-015-389A-130	Sequence 130, App
290	29	76.3	335	4	US-10-230-438-170	Sequence 170, App	363	29	76.3	335	4	US-10-015-519A-130	Sequence 130, App
291	29	76.3	335	4	US-10-232-422-170	Sequence 170, App	364	29	76.3	335	4	US-10-013-915A-130	Sequence 130, App
292	29	76.3	335	4	US-10-223-088-190	Sequence 190, App	365	29	76.3	335	4	US-10-015-394A-130	Sequence 130, App
293	29	76.3	335	4	US-10-223-088-190	Sequence 190, App	366	29	76.3	335	4	US-10-015-390A-130	Sequence 130, App
294	29	76.3	335	4	US-10-223-090-190	Sequence 190, App	367	29	76.3	335	4	US-10-006-746A-130	Sequence 130, App
295	29	76.3	335	4	US-10-219-070-170	Sequence 170, App	368	29	76.3	335	4	US-10-226-254A-130	Sequence 130, App
296	29	76.3	335	4	US-10-219-472-170	Sequence 170, App	369	29	76.3	335	4	US-10-305-654-190	Sequence 190, App
297	29	76.3	335	4	US-10-219-527-170	Sequence 170, App	370	29	76.3	335	4	US-10-011-795A-130	Sequence 130, App
298	29	76.3	335	4	US-10-227-877-170	Sequence 170, App	371	29	76.3	335	4	US-10-232-826-170	Sequence 170, App
299	29	76.3	335	4	US-10-223-087-190	Sequence 190, App	372	29	76.3	335	4	US-10-012-231A-130	Sequence 130, App
300	29	76.3	335	4	US-10-011-692A-130	Sequence 130, App	373	29	76.3	335	4	US-10-230-130-170	Sequence 170, App
301	29	76.3	335	4	US-10-006-768A-130	Sequence 130, App	374	29	76.3	335	4	US-10-616-263-4	Sequence 4, Appli
302	29	76.3	335	4	US-10-017-610A-130	Sequence 130, App	375	29	76.3	335	4	US-10-081-056-190	Sequence 190, App
303	29	76.3	335	4	US-10-006-063A-130	Sequence 130, App	376	29	76.3	335	4	US-10-219-535-170	Sequence 170, App
304	29	76.3	335	4	US-10-020-063A-130	Sequence 130, App	377	29	76.3	335	4	US-10-232-826-170	Sequence 170, App
305	29	76.3	335	4	US-10-223-083-190	Sequence 190, App	378	29	76.3	335	4	US-10-015-395A-130	Sequence 130, App
306	29	76.3	335	4	US-10-216-166-170	Sequence 170, App	379	29	76.3	335	4	US-10-119-480-170	Sequence 170, App
307	29	76.3	335	4	US-10-218-612-170	Sequence 170, App	380	29	76.3	335	5	US-10-219-477-170	Sequence 170, App
308	29	76.3	335	4	US-10-015-391A-130	Sequence 130, App	381	29	76.3	335	5	US-10-012-149A-130	Sequence 130, App
309	29	76.3	335	4	US-10-223-089-190	Sequence 190, App	382	29	76.3	335	5	US-10-233-134-170	Sequence 170, App
310	29	76.3	335	4	US-10-017-407A-130	Sequence 130, App	383	29	76.3	335	5	US-10-219-062-170	Sequence 170, App
311	29	76.3	335	4	US-10-011-833A-130	Sequence 130, App	384	29	76.3	335	5	US-10-219-064-170	Sequence 170, App
312	29	76.3	335	4	US-10-006-041A-130	Sequence 130, App	385	29	76.3	335	5	US-10-218-784-170	Sequence 170, App
313	29	76.3	335	4	US-10-015-822A-130	Sequence 130, App	386	29	76.3	335	5	US-10-219-061-170	Sequence 170, App
314	29	76.3	335	4	US-10-015-827A-130	Sequence 130, App	387	29	76.3	335	6	US-11-025-607-130	Sequence 130, App
315	29	76.3	335	4	US-10-006-130A-130	Sequence 130, App	388	29	76.3	335	5	US-10-739-930-10040	Sequence 10040, A
316	29	76.3	335	4	US-10-216-163-170	Sequence 170, App	389	29	76.3	335	6	US-11-096-568A-25553	Sequence 25553, A
317	29	76.3	335	4	US-10-006-172A-130	Sequence 130, App	390	29	76.3	336	6	US-11-045-004-552	Sequence 552, A
318	29	76.3	335	4	US-10-017-253A-130	Sequence 130, App	391	29	76.3	337	4	US-10-767-701-46915	Sequence 46915, A
319	29	76.3	335	4	US-10-015-392A-130	Sequence 130, App	392	29	76.3	337	4	US-10-425-115-246895	Sequence 246895, A

393	29	76.3	357	4	US-10-425-115-246897
394	29	76.3	357	4	US-10-425-115-246902
395	29	76.3	357	4	US-10-425-115-246905
396	29	76.3	357	4	US-10-425-115-246951
397	29	76.3	357	4	US-10-425-115-246541
398	29	76.3	357	5	US-10-425-115-246543
399	29	76.3	357	5	US-10-425-115-246543
400	29	76.3	360	3	US-09-734-569-168
401	29	76.3	361	4	US-10-437-963-152533
402	29	76.3	369	4	US-10-425-115-300452
403	29	76.3	381	4	US-10-425-115-246901
404	29	76.3	392	4	US-10-425-115-246906
405	29	76.3	398	4	US-10-425-115-338541
406	29	76.3	407	4	US-10-425-114-70879
407	29	76.3	413	4	US-10-425-115-338543
408	29	76.3	414	4	US-10-425-114-65555
409	29	76.3	418	6	US-11-096-568A-25551
410	29	76.3	428	4	US-10-425-114-65664
411	29	76.3	430	4	US-10-437-963-104713
412	29	76.3	437	4	US-10-467-824-22
413	29	76.3	445	4	US-10-360-101-266
414	29	76.3	449	5	US-10-425-115-338548
415	29	76.3	463	6	US-11-188-298-12963
416	29	76.3	500	6	US-11-079-463-5773
417	29	76.3	536	4	US-10-424-539-180356
418	29	76.3	545	4	US-10-369-493-4055
419	29	76.3	585	4	US-10-369-493-9765
420	29	76.3	598	4	US-10-369-493-18831
421	29	76.3	613	4	US-10-425-115-274957
422	29	76.3	689	5	US-10-450-763-30400
423	29	76.3	767	6	US-11-188-298-12159
424	29	76.3	806	4	US-10-363-616-322
425	29	76.3	806	4	US-10-408-765A-833
426	29	76.3	807	5	US-10-733-923-12960
427	29	76.3	810	5	US-10-733-923-12960
428	29	76.3	810	5	US-10-755-149-5159
429	29	76.3	878	4	US-10-437-963-164591
430	29	76.3	929	5	US-10-733-923-12929
431	29	76.3	933	5	US-10-733-923-12866
432	29	76.3	933	5	US-10-733-923-12962
433	29	76.3	933	5	US-10-733-923-12963
434	29	76.3	933	5	US-10-733-923-12966
435	29	76.3	949	5	US-10-733-923-12930
436	29	76.3	949	5	US-10-733-923-12931
437	29	76.3	954	5	US-10-733-923-12961
438	29	76.3	954	5	US-10-733-923-12966
439	29	76.3	954	5	US-10-733-923-12969
440	29	76.3	1065	4	US-10-282-122A-64932
441	29	76.3	1456	4	US-10-437-963-194566
442	29	76.3	1818	4	US-10-032-585-7616
443	29	76.3	1878	6	US-11-097-143-12348
444	29	76.3	2005	5	US-10-450-763-54823
445	29	76.3	4101	4	US-10-369-493-6509
446	29	76.3	4101	4	US-10-369-493-6510
447	29	76.3	4928	4	US-10-328-148A-5
448	29	76.3	25	4	US-10-029-386-27689
449	29	76.3	34	4	US-10-425-115-205847
450	29	76.3	39	4	US-10-424-599-225682
451	29	76.3	40	4	US-10-424-599-282017
452	29	76.3	41	4	US-10-425-115-306114
453	29	76.3	43	4	US-10-424-599-181514
454	29	76.3	47	4	US-10-424-599-157183
455	29	76.3	45	4	US-10-425-115-550220
456	29	76.3	52	4	US-10-425-114-50413
457	29	76.3	53	4	US-10-767-701-48245
458	29	76.3	54	4	US-10-424-599-146904
459	29	76.3	56	4	US-10-425-115-292216
460	29	76.3	60	3	US-09-864-761-48786
461	29	76.3	61	4	US-10-106-698-8347
462	29	76.3	62	4	US-10-425-115-306067
463	29	76.3	63	5	US-10-450-763-39220
464	29	76.3	66	4	US-10-425-115-284754
465	29	76.3	69	4	US-10-424-599-157998

Sequence 246897,	466	28	73.7	70	4	US-10-424-599-174534	Sequence 174534,
Sequence 246902,	467	28	73.7	70	4	US-10-425-115-307311	Sequence 307311,
Sequence 246905,	468	28	73.7	71	4	US-10-335-977-9307	Sequence 9307, Ap
Sequence 346541,	469	28	73.7	75	3	US-09-864-761-45039	Sequence 45039, A
Sequence 346544,	470	28	73.7	75	4	US-10-335-977-9308	Sequence 9308, Ap
Sequence 10029, A	471	28	73.7	77	4	US-10-425-115-299471	Sequence 299471, Ap
Sequence 168, App	472	28	73.7	80	3	US-09-889-890-194	Sequence 194, App
Sequence 132533,	473	28	73.7	82	4	US-10-425-115-266020	Sequence 266020, A
Sequence 300450,	474	28	73.7	83	4	US-10-425-115-266021	Sequence 266021,
Sequence 246901,	475	28	73.7	86	4	US-10-424-599-249255	Sequence 249255,
Sequence 338541,	476	28	73.7	89	4	US-10-002-344A-168	Sequence 168, App
Sequence 338544,	477	28	73.7	89	5	US-10-958-862-168	Sequence 168, App
Sequence 70879, A	478	28	73.7	90	4	US-10-424-599-218746	Sequence 218746,
Sequence 338543,	479	28	73.7	98	4	US-10-115-879-24	Sequence 24, Appl
Sequence 65555, A	480	28	73.7	100	4	US-10-029-386-28417	Sequence 28417, A
Sequence 25561, A	481	28	73.7	103	3	US-09-833-245-1395	Sequence 1395, Ap
Sequence 65664, A	482	28	73.7	103	3	US-09-833-245-1396	Sequence 1396, Ap
Sequence 104713,	483	28	73.7	103	4	US-10-094-749-2935	Sequence 2935, Ap
Sequence 22, Appl	484	28	73.7	103	6	US-11-264-096-1396	Sequence 1396, Ap
Sequence 226, App	485	28	73.7	105	6	US-10-437-963-124699	Sequence 124699, A
Sequence 338548,	486	28	73.7	105	6	US-11-096-568A-23057	Sequence 23057, A
Sequence 314, App	487	28	73.7	108	3	US-09-867-550-278	Sequence 278, App
Sequence 12963, A	488	28	73.7	109	4	US-10-437-963-16810	Sequence 16810,
Sequence 5773, Ap	489	28	73.7	114	4	US-10-424-599-236283	Sequence 236283,
Sequence 180356,	490	28	73.7	116	4	US-10-424-599-230737	Sequence 230737,
Sequence 4055, Ap	491	28	73.7	122	4	US-10-424-599-200697	Sequence 200697,
Sequence 160356,	492	28	73.7	124	6	US-11-096-568A-23056	Sequence 23056, A
Sequence 274957,	493	28	73.7	126	4	US-10-437-963-200586	Sequence 200586,
Sequence 30400, A	494	28	73.7	127	4	US-10-767-701-33318	Sequence 33318, A
Sequence 32159, A	495	28	73.7	127	4	US-10-425-115-304184	Sequence 304184,
Sequence 833, App	496	28	73.7	140	4	US-10-425-115-294547	Sequence 294547,
Sequence 12960, A	497	28	73.7	141	4	US-10-767-701-39126	Sequence 39126, A
Sequence 1147, Ap	498	28	73.7	143	4	US-10-425-115-353247	Sequence 353247,
Sequence 15159, Ap	499	28	73.7	144	4	US-10-424-599-201869	Sequence 201869,
Sequence 164591,	500	28	73.7				

ALIGNMENTS

```
RESULT 1
US-10-541-343-3
; Sequence 3, Application US/10541343
; Publication No. US20060052299A1
; GENERAL INFORMATION:
; APPLICANT: Chau, Raymond Ming Wah
; APPLICANT: Ko, Pui-Yuk Dorothy
; TITLE OF INVENTION: MNTF Peptides and Compositions and Methods of Use
; FILE REFERENCE: 2001-103US
; CURRENT APPLICATION NUMBER: US/10/541,343
; PRIOR FILING DATE: 2005-07-05
; PRIOR APPLICATION NUMBER: PCT/US2004/001468
; PRIOR FILING DATE: 2004-01-21
; PRIOR APPLICATION NUMBER: 60/441,722
; PRIOR FILING DATE: 2003-01-21
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
US-10-541-343-3

Query Match      100.0%; Score 38; DB 5; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.9e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Oy	1	WMLSAFS	7
Db	1	WMLSAFS	7

```
RESULT 2
US-10-541-343-6
; Sequence 6, Application US/10541343
; Publication No. US2006052299A1
; GENERAL INFORMATION:
; APPLICANT: Chau, Raymond Ming Wah
; APPLICANT: Ko, Pui-Yuk Dorothy
; TITLE OF INVENTION: MNTF Peptides and Compositions and Methods of Use
; FILE REFERENCE: 2001-103US
; CURRENT APPLICATION NUMBER: US/10/541,343
; CURRENT FILING DATE: 2005-07-05
; PRIOR APPLICATION NUMBER: PCT/US2004/001468
; PRIOR FILING DATE: 2004-01-21
; PRIOR APPLICATION NUMBER: 60/441,722
; PRIOR FILING DATE: 2003-01-21
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 6
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
US-10-541-343-6

Query Match      100.0%; Score 38; DB 5; Length 13;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WMLSAFS 7
        |||||
Db      2 WMLSAFS 8

RESULT 3
US-09-989-481-4
; Sequence 4, Application US/09989481
; Publication No. US20020086831A1
; GENERAL INFORMATION:
; APPLICANT: Chau, Raymond M.W.
; TITLE OF INVENTION: Isolation and Use of Motoneuronotropic Factors
; FILE REFERENCE: 12592-4
; CURRENT APPLICATION NUMBER: US/09/989,481
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: US 09/633,447
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 08/9928862
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: US 08/751225
; PRIOR FILING DATE: 1996-11-15
; PRIOR APPLICATION NUMBER: US 60/026792
; PRIOR FILING DATE: 1996-09-27
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-989-481-4

Query Match      100.0%; Score 38; DB 3; Length 33;
Best Local Similarity 100.0%; Pred. No. 6; 8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WMLSAFS 7
        |||||
Db      12 WMLSAFS 18

RESULT 4
US-10-858-545-4
; Sequence 4, Application US/10858545
```

```
; Publication No. US20040224894A1
; GENERAL INFORMATION:
; APPLICANT: Chau, Raymond M.W.
; TITLE OF INVENTION: Isolation and Use of Motoneuronotropic Factors
; FILE REFERENCE: 12592-3
; CURRENT APPLICATION NUMBER: US/10/858,545
; CURRENT FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US/09/592,018
; PRIOR FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: US 08/9928862
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: US 08/751225
; PRIOR FILING DATE: 1996-11-15
; PRIOR APPLICATION NUMBER: US 60/026792
; PRIOR FILING DATE: 1996-09-27
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-858-545-4

Query Match      100.0%; Score 38; DB 5; Length 33;
Best Local Similarity 100.0%; Pred. No. 6; 8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WMLSAFS 7
        |||||
Db      12 WMLSAFS 18

RESULT 5
US-10-858-286-4
; Sequence 4, Application US/10858286
; Publication No. US20040242486A1
; GENERAL INFORMATION:
; APPLICANT: Chau, Raymond M.W.
; TITLE OF INVENTION: Isolation and Use of Motoneuronotropic Factors
; FILE REFERENCE: 12592-3
; CURRENT APPLICATION NUMBER: US/10/858,286
; CURRENT FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US/09/592,018
; PRIOR FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: US 08/9928862
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: US 08/751225
; PRIOR FILING DATE: 1996-11-15
; PRIOR APPLICATION NUMBER: US 60/026792
; PRIOR FILING DATE: 1996-09-27
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-858-286-4

Query Match      100.0%; Score 38; DB 5; Length 33;
Best Local Similarity 100.0%; Pred. No. 6; 8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WMLSAFS 7
        |||||
Db      12 WMLSAFS 18

RESULT 6
US-10-858-543-4
; Sequence 4, Application US/10858543
; Publication No. US2004026691A1
; GENERAL INFORMATION:
; APPLICANT: Chau, Raymond M.W.
```

;; TITLE OF INVENTION: Isolation and Use of Motoneuronotropic Factors
;; FILE REFERENCE: 12592-3
;; CURRENT APPLICATION NUMBER: US/10/858,543
;; CURRENT FILING DATE: 2004-06-01
;; PRIOR APPLICATION NUMBER: US/09/592,018
;; PRIOR FILING DATE: 2000-06-12
;; PRIOR APPLICATION NUMBER: US 08/9928862
;; PRIOR FILING DATE: 1997-09-12
;; PRIOR APPLICATION NUMBER: US 08/751225
;; PRIOR FILING DATE: 1996-11-15
;; PRIOR APPLICATION NUMBER: US 60/026792
;; PRIOR FILING DATE: 1996-09-27
;; NUMBER OF SEQ ID NOS: 5
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 4
;; LENGTH: 33
;; TYPE: PRT
;; ORGANISM: Homo Sapiens
US-10-858-543-4

Query Match 100.0%; Score 38; DB 5; Length 33;
Best Local Similarity 100.0%; Pred. No. 6.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
Db 12 WMLSAFS 18

RESULT 7
US-10-858-144-4
;; Sequence 4, Application US/10858144
;; Publication No. US20060025565A1
;; GENERAL INFORMATION:
;; APPLICANT: Chau, Raymond M.W.
;; TITLE OF INVENTION: Isolation and Use of Motoneuronotropic Factors
;; FILE REFERENCE: 12592-3
;; CURRENT APPLICATION NUMBER: US/10/858,144
;; CURRENT FILING DATE: 2004-06-01
;; PRIOR APPLICATION NUMBER: US/09/592,018
;; PRIOR FILING DATE: 2000-06-12
;; PRIOR APPLICATION NUMBER: US 08/9928862
;; PRIOR FILING DATE: 1997-09-12
;; PRIOR APPLICATION NUMBER: US 08/751225
;; PRIOR FILING DATE: 1996-11-15
;; PRIOR APPLICATION NUMBER: US 60/026792
;; PRIOR FILING DATE: 1996-09-27
;; NUMBER OF SEQ ID NOS: 5
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 4
;; LENGTH: 33
;; TYPE: PRT
;; ORGANISM: Homo Sapiens
US-10-858-144-4

Query Match 100.0%; Score 38; DB 5; Length 33;
Best Local Similarity 100.0%; Pred. No. 6.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
Db 12 WMLSAFS 18

RESULT 8
US-10-541-343-1
;; Sequence 1, Application US/10541343
;; Publication No. US20060052299A1
;; GENERAL INFORMATION:
;; APPLICANT: Chau, Raymond Ming Wah
;; APPLICANT: Ko, Pui-Yuk Dorothy
;; TITLE OF INVENTION: WNTF Peptides and Compositions and Methods of Use
;; FILE REFERENCE: 2001-1030US

;; CURRENT APPLICATION NUMBER: US/10/541,343
;; CURRENT FILING DATE: 2005-07-05
;; PRIOR APPLICATION NUMBER: PCT/US2004/001468
;; PRIOR FILING DATE: 2004-01-21
;; PRIOR APPLICATION NUMBER: 60/441,722
;; PRIOR FILING DATE: 2003-01-21
;; NUMBER OF SEQ ID NOS: 7
;; SOFTWARE: PatentIn version 3.3
;; SEQ ID NO 1
;; LENGTH: 33
;; TYPE: PRT
;; ORGANISM: Artificial
;; FEATURE:
;; OTHER INFORMATION: Synthetic peptide
US-10-541-343-1

Query Match 100.0%; Score 38; DB 5; Length 33;
Best Local Similarity 100.0%; Pred. No. 6.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
Db 12 WMLSAFS 18

RESULT 9
US-11-097-143-16893
;; Sequence 16893, Application US/11097143
;; Publication No. US2005020858A1
;; GENERAL INFORMATION:
;; APPLICANT: Venter, J. Craig
;; APPLICANT: et al.
;; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
;; ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
;; FILE REFERENCE: C0000728
;; TITLE OF INVENTION: DROSOPHILA GENES.
;; CURRENT APPLICATION NUMBER: US/11/097,143
;; CURRENT FILING DATE: 2005-04-04
;; PRIOR APPLICATION NUMBER: 60/157,832
;; PRIOR FILING DATE: 1999-10-05
;; PRIOR APPLICATION NUMBER: 60/160,191
;; PRIOR FILING DATE: 1999-10-19
;; PRIOR APPLICATION NUMBER: 60/161,932
;; PRIOR FILING DATE: 1999-10-28
;; PRIOR APPLICATION NUMBER: 60/164,769
;; PRIOR FILING DATE: 1999-11-12
;; PRIOR APPLICATION NUMBER: 60/173,383
;; PRIOR FILING DATE: 1999-12-28
;; PRIOR APPLICATION NUMBER: 60/175,693
;; PRIOR FILING DATE: 2000-01-12
;; PRIOR APPLICATION NUMBER: 60/184,831
;; PRIOR FILING DATE: 2000-02-24
;; PRIOR APPLICATION NUMBER: 60/191,637
;; PRIOR FILING DATE: 2000-03-23
;; NUMBER OF SEQ ID NOS: 43008
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 16893
;; LENGTH: 1760
;; TYPE: PRT
;; ORGANISM: DROSOPHILA
US-11-097-143-16893

Query Match 92.1%; Score 35; DB 6; Length 1760;
Best Local Similarity 85.7%; Pred. No. 8.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
Db 966 WMLSAFS 972

RESULT 10
US-10-424-599-187845

; Sequence 187845, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ. ID NOS: 285684
; SEQ. ID NO 187845
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_140637C.1.pep
; US-10-424-599-187845

Query Match 89.5%; Score 34; DB 4; Length 103;
Best Local Similarity 85.7%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSPS 7
Db 29 WMLGAPS 35

RESULT 11

US-10-124-557-18
; Sequence 18, Application US/10124557
; Publication No. US20020137894A1
; GENERAL INFORMATION:

APPLICANT: Turner, Katherine
Clark, Stephen C.
Jacobs, Kenneth C.
Hewick, Rodney M.
Gesner, Thomas G.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/124,557
FILING DATE: 16-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991

APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Cserr, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 876-1170

TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 46 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-10-124-557-18

Query Match 86.8%; Score 33; DB 4; Length 46;
Best Local Similarity 85.7%; Pred. No. 78;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSPS 7
Db 5 WMLSPS 11

RESULT 12

US-11-169-232-18
; Sequence 18, Application US/11169232
; Publication No. US20060025570A1
; GENERAL INFORMATION:

APPLICANT: Turner, Katherine
Clark, Stephen C.
Jacobs, Kenneth
Hewick, Rodney M.
Gesner, Thomas G.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/11/169,232
FILING DATE: 28-Jun-2005
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/10/124,557
FILING DATE: 16-Apr-2002

APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Cserr, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 876-1170
TELEFAX: (617) 876-5851

INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-11-169-232-18

Query Match 86.8%; Score 33; DB 6; Length 46;
Best Local Similarity 85.7%; Pred. No. 78;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSPFS 7
Db 5 WMLSPFS 11

RESULT 13

US-09-925-301-1452
Sequence 1452, Application US/09925301
Patent No. US20020052308A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA106
CURRENT APPLICATION NUMBER: US/09/925,301
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05882
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1694
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1452
LENGTH: 61
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (6)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-301-1452

Query Match 86.8%; Score 33; DB 3; Length 61;
Best Local Similarity 85.7%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSPFS 7
Db 37 WMLSPFS 43

RESULT 14

US-10-124-557-72
Sequence 72, Application US/10124557
Publication No. US20020137894A1
GENERAL INFORMATION:
APPLICANT: Turner, Katherine
Clark, Stephen C.
Jacobs, Kenneth
Hewick, Rodney M.
Gesner, Thomas G.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/124,557
FILING DATE: 16-Apr-2002
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Cserr, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 876-1170
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 72:
SEQUENCE CHARACTERISTICS:
LENGTH: 237 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 72:
US-10-124-557-72

Query Match 86.8%; Score 33; DB 4; Length 237;
Best Local Similarity 85.7%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSPFS 7
Db 111 WMLSPFS 117

RESULT 15

US-11-169-232-72
Sequence 72, Application US/11169232
Publication No. US20060025570A1
GENERAL INFORMATION:
APPLICANT: Turner, Katherine
Clark, Stephen C.
Jacobs, Kenneth
Hewick, Rodney M.
Gesner, Thomas G.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/11/169,232
FILING DATE: 28-Jun-2005
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/10/124,557
FILING DATE: 16-Apr-2002
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989

ATTORNEY/AGENT INFORMATION:
NAME: Cserr, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 876-1170
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 72:
SEQUENCE CHARACTERISTICS:
LENGTH: 237 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 72:
US-11-169-232-72

Query Match 86.8%; Score 33; DB 6; Length 237;
Best Local Similarity 85.7%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSPS 7
Db 111 WMLSPS 117

RESULT 16
US-10-124-557-64
Sequence 64, Application US/10124557
Publication No. US20020137894A1
GENERAL INFORMATION:
APPLICANT: Turner, Katherine
Clark, Stephen C.
Jacobs, Kenneth
Hewick, Rodney M.
Gesner, Thomas G.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:
ADDRESSER: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/124,557
FILING DATE: 16-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Cserr, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 876-1170
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 64:
SEQUENCE CHARACTERISTICS:
LENGTH: 372 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 64:
US-10-124-557-64

Query Match 86.8%; Score 33; DB 4; Length 372;
Best Local Similarity 85.7%; Pred. No. 5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSPS 7
Db 139 WMLSPS 145

RESULT 17
US-11-169-232-64
Sequence 64, Application US/11169232
Publication No. US20060025570A1
GENERAL INFORMATION:
APPLICANT: Turner, Katherine
Clark, Stephen C.
Jacobs, Kenneth
Hewick, Rodney M.
Gesner, Thomas G.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:
ADDRESSER: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/11/169,232
FILING DATE: 28-Jun-2005
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/10/124,557
FILING DATE: 16-Apr-2002
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Cserr, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 876-1170
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 64:
SEQUENCE CHARACTERISTICS:
LENGTH: 372 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 64:
US-11-169-232-64

Query Match 86.8%; Score 33; DB 6; Length 372;
Best Local Similarity 85.7%; Pred. No. 5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
Db 139 WMLSPFS 145

RESULT 18
US-10-124-557-68
; Sequence 68, Application US/10124557
; Publication No. US20020137894A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Gesner, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cserr, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)876-1170
; TELEFAX: (617)876-5851
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 422 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 68:
US-10-124-557-68

Query Match 86.8%; Score 33; DB 4; Length 422;
Best Local Similarity 85.7%; Pred. No. 5.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
Db 189 WMLSPFS 195

RESULT 19
US-11-169-232-68
; Sequence 68, Application US/11169232
; Publication No. US20060025570A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine

Clark, Stephen C.
Jacobs, Kenneth
Hewick, Rodney M.
Gesner, Thomas G.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/11/169,232
FILING DATE: 28-Jun-2005
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/10/124,557
FILING DATE: 16-Apr-2002
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Cserr, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)876-1170
TELEFAX: (617)876-5851
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 422 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 68:
US-11-169-232-68

Query Match 86.8%; Score 33; DB 6; Length 422;
Best Local Similarity 85.7%; Pred. No. 5.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
Db 189 WMLSPFS 195

RESULT 20
US-10-124-557-66
; Sequence 66, Application US/10124557
; Publication No. US20020137894A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Gesner, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive

CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/124,557
FILING DATE: 16-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Cserr, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 876-1170
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 66:
SEQUENCE CHARACTERISTICS:
LENGTH: 423 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 66:
US-10-124-557-66
Query Match 86.8%; Score 33; DB 4; Length 423;
Best Local Similarity 85.7%; Pred. No. 5.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 WMLSPS 7
DB 190 WMLSPS 196
RESULT 21
US-11-169-232-66
Sequence 66, Application US/11169232
Publication No. US2006025570A1
GENERAL INFORMATION:
APPLICANT: Turner, Katherine
Clark, Stephen C.
Jacobs, Kenneth C.
Hewick, Rodney M.
Gesner, Thomas G.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/11/169,232

FILING DATE: 28-Jun-2005
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/10/124,557
FILING DATE: 16-Apr-2002
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Cserr, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 876-1170
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 66:
SEQUENCE CHARACTERISTICS:
LENGTH: 423 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 66:
US-11-169-232-66
Query Match 86.8%; Score 33; DB 6; Length 423;
Best Local Similarity 85.7%; Pred. No. 5.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 WMLSPS 7
DB 190 WMLSPS 196
RESULT 22
US-10-124-557-54
Sequence 54, Application US/10124557
Publication No. US20020137894A1
GENERAL INFORMATION:
APPLICANT: Turner, Katherine
Clark, Stephen C.
Jacobs, Kenneth C.
Hewick, Rodney M.
Gesner, Thomas G.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/124,557
FILING DATE: 16-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989

APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Cseert, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)876-1170
TELEFAX: (617)876-5851
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 463 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 54:
US-10-124-557-54

Query Match 86.8%; Score 33; DB 4; Length 463;
Best Local Similarity 85.7%; Pred. No. 6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
Db 230 WMLSPFS 236

RESULT 23
US-11-169-232-54
Sequence 54, Application US/11169232
Publication No. US20060025570A1
GENERAL INFORMATION:
APPLICANT: Turner, Katherine
Clark, Stephen C.
Jacobs, Kenneth
Hewick, Rodney M.
Gesner, Thomas G.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/11/169,232
FILING DATE: 28-Jun-2005
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/10/124,557
FILING DATE: 16-Apr-2002
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JUN-1991
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Cseert, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)876-1170
TELEFAX: (617)876-5851

INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 463 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 54:
US-11-169-232-54

Query Match 86.8%; Score 33; DB 6; Length 463;
Best Local Similarity 85.7%; Pred. No. 6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
Db 230 WMLSPFS 236

RESULT 24
US-10-038-694-3
Sequence 3, Application US/10038694
Publication No. US20030180948A1
GENERAL INFORMATION:
APPLICANT: Dixon, Eric
APPLICANT: Hutchins, Jeff T.
APPLICANT: Kuettnier, Klaus E.
APPLICANT: Schmid, Thomas M.
APPLICANT: Schumacher, Barbara L.
APPLICANT: Su, Jui-Ian
TITLE OF INVENTION: SUPERFICIAL ZONE PROTEIN AND METHODS OF
FILE OF INVENTION: MAKING AND USING SAME
FILE REFERENCE: 07083,000805
CURRENT APPLICATION NUMBER: US/10/038,694
CURRENT FILING DATE: 2001-12-31
PRIOR APPLICATION NUMBER: 60/258,920
PRIOR FILING DATE: 2000-12-29
NUMBER OF SEQ ID NOS: 11
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 538
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence; note =
US-10-038-694-3

Query Match 86.8%; Score 33; DB 4; Length 538;
Best Local Similarity 85.7%; Pred. No. 6.9e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
Db 305 WMLSPFS 311

RESULT 25
US-09-802-207-27
Sequence 27, Application US/09802207
Publication No. US20020086824A1
GENERAL INFORMATION:
APPLICANT: Warman, Matthew
APPLICANT: Carpen, John
APPLICANT: Trent, Jeffrey
APPLICANT: Marcelino, Jose
TITLE OF INVENTION: Novel Methods and Reagents for the Treatment of Osteoarthritis
FILE REFERENCE: Case-06212
CURRENT APPLICATION NUMBER: US/09/802,207
CURRENT FILING DATE: 2001-08-29
PRIOR APPLICATION NUMBER: 09/619,175
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 60/145,328
PRIOR FILING DATE: 1999-07-23

; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 27
; LENGTH: 792
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: NON CONS
; LOCATION: (267)..(268)
; NAME/KEY: NON CONS
; LOCATION: (321)..(322)
US-09-802-207-27

Query Match 86.8%; Score 33; DB 3; Length 792;
Best Local Similarity 85.7%; Pred. No. 9.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSPFS 7
Db 559 WMLSPFS 565

RESULT 26
US-10-425-115-337015
; Sequence 337015, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 337015
; LENGTH: 1046
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1046)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_70527C.1.pcp
US-10-425-115-337015

Query Match 86.8%; Score 33; DB 4; Length 1046;
Best Local Similarity 71.4%; Pred. No. 1.2e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 WMLSPFS 7
Db 890 WMLSPFS 896

RESULT 27
US-10-124-557-44
; Sequence 44, Application US/10124557
; Publication No. US20020137894A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; APPLICANT: Clark, Stephen C.
; APPLICANT: Jacobs, Kenneth
; APPLICANT: Hewick, Rodney M.
; APPLICANT: Gesner, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive

; City: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140

; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cseri, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 876-1170
; TELEFAX: (617) 876-5851

; INFORMATION FOR SEQ ID NO: 44:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1270 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 44:
US-10-124-557-44

Query Match 86.8%; Score 33; DB 4; Length 1270;
Best Local Similarity 85.7%; Pred. No. 1.5e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSPFS 7
Db 1037 WMLSPFS 1043

RESULT 28
US-11-169-232-44
; Sequence 44, Application US/11169232
; Publication No. US2006002570A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; APPLICANT: Clark, Stephen C.
; APPLICANT: Jacobs, Kenneth
; APPLICANT: Hewick, Rodney M.
; APPLICANT: Gesner, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/11/169,232

FILING DATE: 28-Jun-2005
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/10/124,557
FILING DATE: 16-Apr-2002
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-Jan-1991
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-Jun-1990
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-Dec-1989
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-Aug-1989
ATTORNEY/AGENT INFORMATION:
NAME: Cseerr, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)876-1170
TELEFAX: (617)876-5851
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 1270 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 44:
US-11-169-232-44

Query Match 86.8%; Score 33; DB 6; Length 1270;
Best Local Similarity 85.7%; Pred. No. 1.5e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 WMLSPFS 7
Db 1037 WMLSPFS 1043

RESULT 29
US-10-124-557-42
Sequence 42, Application US/10124557
Publication No. US20020137894A1
GENERAL INFORMATION:
APPLICANT: Turner, Katherine
Clark, Stephen C.
Jacobs, Kenneth
Gesner, Rodney M.
Hewick, Thomas G.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/124,557
FILING DATE: 16-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-Jan-1991
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-Jun-1990
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-Dec-1989
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-Aug-1989
ATTORNEY/AGENT INFORMATION:
NAME: Cseerr, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)876-1170
TELEFAX: (617)876-5851

APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-Aug-1989
ATTORNEY/AGENT INFORMATION:
NAME: Cseerr, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)876-1170
TELEFAX: (617)876-5851
INFORMATION FOR SEQ ID NO: 42:
SEQUENCE CHARACTERISTICS:
LENGTH: 1311 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 42:
US-10-124-557-42

Query Match 86.8%; Score 33; DB 4; Length 1311;
Best Local Similarity 85.7%; Pred. No. 1.5e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 WMLSPFS 7
Db 1078 WMLSPFS 1084

RESULT 30
US-11-169-232-42
Sequence 42, Application US/11169232
Publication No. US20060025570A1
GENERAL INFORMATION:
APPLICANT: Turner, Katherine
Clark, Stephen C.
Jacobs, Kenneth
Gesner, Rodney M.
Hewick, Thomas G.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/11/169,232
FILING DATE: 28-Jun-2005
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/10/124,557
FILING DATE: 16-Apr-2002
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-Jan-1991
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-Jun-1990
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-Dec-1989
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-Aug-1989
ATTORNEY/AGENT INFORMATION:
NAME: Cseerr, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)876-1170
TELEFAX: (617)876-5851

INFORMATION FOR SEQ ID NO: 42:
SEQUENCE CHARACTERISTICS:
LENGTH: 1311 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 42:
US-11-169-232-42

Query Match 86.8%; Score 33; DB 6; Length 1311;
Best Local Similarity 85.7%; Pred. No. 1.5e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
DB 1078 WMLSPFS 1084

RESULT 31

US-10-124-557-142
Sequence 142, Application US/10124557
Publication No. US20020137894A1

GENERAL INFORMATION:

APPLICANT: Turner, Katherine
Clark, Stephen C.
Jacobs, Kenneth
Hewick, Rodney M.
Gesner, Thomas G.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/124,557

FILING DATE: 16-Apr-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/643,502

FILING DATE: 18-JAN-1991

APPLICATION NUMBER: US 07/546,114

FILING DATE: 29-JUN-1990

APPLICATION NUMBER: US 07/457,196

FILING DATE: 29-DEC-1989

APPLICATION NUMBER: US 07/390,901

FILING DATE: 08-AUG-1989

ATTORNEY/AGENT INFORMATION:

NAME: Cserr, Luann

REGISTRATION NUMBER: 31,822

REFERENCE/DOCKET NUMBER: GI 5190

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)876-1170

TELEFAX: (617)876-5851

INFORMATION FOR SEQ ID NO: 142:

SEQUENCE CHARACTERISTICS:

LENGTH: 1313 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 142:
US-10-124-557-142

Query Match 86.8%; Score 33; DB 4; Length 1313;
Best Local Similarity 85.7%; Pred. No. 1.5e+03;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
DB 1080 WMLSPFS 1086

RESULT 32

US-11-169-232-142

Sequence 142, Application US/11169232

Publication No. US2006002570A1

GENERAL INFORMATION:

APPLICANT: Turner, Katherine
Clark, Stephen C.
Jacobs, Kenneth
Hewick, Rodney M.
Gesner, Thomas G.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/11/169,232

FILING DATE: 28-Jun-2005

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/10/124,557

FILING DATE: 16-Apr-2002

APPLICATION NUMBER: US 07/643,502

FILING DATE: 18-JAN-1991

APPLICATION NUMBER: US 07/546,114

FILING DATE: 29-JUN-1990

APPLICATION NUMBER: US 07/457,196

FILING DATE: 29-DEC-1989

APPLICATION NUMBER: US 07/390,901

FILING DATE: 08-AUG-1989

ATTORNEY/AGENT INFORMATION:

NAME: Cserr, Luann

REGISTRATION NUMBER: 31,822

REFERENCE/DOCKET NUMBER: GI 5190

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)876-1170

TELEFAX: (617)876-5851

INFORMATION FOR SEQ ID NO: 142:

SEQUENCE CHARACTERISTICS:

LENGTH: 1313 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 142:
US-11-169-232-142

Query Match 86.8%; Score 33; DB 6; Length 1313;
Best Local Similarity 85.7%; Pred. No. 1.5e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
DB 1080 WMLSPFS 1086

RESULT 33

US-10-124-557-50

Sequence 50, Application US/10124557
Publication No. US20020137894A1
GENERAL INFORMATION:
APPLICANT: Turner, Katherine
Clark, Stephen C.
Jacobs, Kenneth
Hewick, Rodney M.
Gesner, Thomas G.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/124,557
FILING DATE: 16-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Geert, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 876-1170
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 1314 amino acids
TYPE: amino acid
MOLECULE TYPE: protein
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 50:
US-10-124-557-50

Query Match 86.8%; Score 33; DB 4; Length 1314;
Best Local Similarity 85.7%; Pred. No. 1.5e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
Db 1081 WMLSPFS 1087

RESULT 34
US-11-169-232-50
Sequence 50, Application US/11169232
Publication No. US20060025570A1
GENERAL INFORMATION:
APPLICANT: Turner, Katherine
Clark, Stephen C.
Jacobs, Kenneth
Hewick, Rodney M.
Gesner, Thomas G.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESSES:
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/11/169,232
FILING DATE: 28-Jun-2005
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/10/124,557
FILING DATE: 16-Apr-2002
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Geert, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 876-1170
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 1314 amino acids
TYPE: amino acid
MOLECULE TYPE: protein
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 50:
US-11-169-232-50

Query Match 86.8%; Score 33; DB 6; Length 1314;
Best Local Similarity 85.7%; Pred. No. 1.5e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
Db 1081 WMLSPFS 1087

RESULT 35
US-10-124-557-46
Sequence 46, Application US/10124557
Publication No. US20020137894A1
GENERAL INFORMATION:
APPLICANT: Turner, Katherine
Clark, Stephen C.
Jacobs, Kenneth
Hewick, Rodney M.
Gesner, Thomas G.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/124,557
FILING DATE: 16-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Cserr, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 876-1170
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 1320 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 46:
US-10-124-557-46

Query Match 86.8%; Score 33; DB 4; Length 1320;
Best Local Similarity 85.7%; Pred. No. 1.5e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSPFS 7
Db 1087 WMLSPFS 1093

RESULT 36
US-10-124-557-60
Sequence 60, Application US/10124557
Publication No. US20020137894A1
GENERAL INFORMATION:
APPLICANT: Turner, Katherine
Clark, Stephen C.
Jacobs, Kenneth
Hewick, Rodney M.
Gesner, Thomas G.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/124,557
FILING DATE: 16-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990

APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Cserr, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 876-1170
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 60:
SEQUENCE CHARACTERISTICS:
LENGTH: 1320 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 60:
US-10-124-557-60

Query Match 86.8%; Score 33; DB 4; Length 1320;
Best Local Similarity 85.7%; Pred. No. 1.5e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSPFS 7
Db 1087 WMLSPFS 1093

RESULT 37
US-10-717-665-58
Sequence 58, Application US/10717665
Publication No. US20050106579A1
GENERAL INFORMATION:
APPLICANT: Origene Technologies, Inc
TITLE OF INVENTION: Regulated Angiogenesis Genes and Polypeptides
FILE REFERENCE: IU 103 R1
CURRENT APPLICATION NUMBER: US/10/717,665
CURRENT FILING DATE: 2003-11-21
PRIOR APPLICATION NUMBER: US/10/164,595
PRIOR FILING DATE: 2002-06-10
NUMBER OF SEQ ID NOS: 80
SOFTWARE: PatentIn version 3.1
SEQ ID NO 58
LENGTH: 1320
TYPE: PRT
ORGANISM: Homo sapiens
US-10-717-665-58

Query Match 86.8%; Score 33; DB 5; Length 1320;
Best Local Similarity 85.7%; Pred. No. 1.5e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSPFS 7
Db 1087 WMLSPFS 1093

RESULT 38
US-11-169-232-46
Sequence 46, Application US/11169232
Publication No. US20060025570A1
GENERAL INFORMATION:
APPLICANT: Turner, Katherine
Clark, Stephen C.
Jacobs, Kenneth
Hewick, Rodney M.
Gesner, Thomas G.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive

```

; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/11/169,232
; FILING DATE: 28-Jun-2005
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-Jan-1991
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-Jun-1990
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cserr, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)876-1170
; TELEFAX: (617)876-5851
;
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1320 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 46:
;
; US-11-169-232-46
;
Query Match      86.8%; Score 33; DB 6; Length 1320;
Best Local Similarity 85.7%; Pred. No. 1.5e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
;
QY      1 WMLSAFS 7
Db      1087 WMLSPFS 1093
;
RESULT 39
US-11-169-232-60
; Sequence 60, Application US/11169232
; Publication No. US20060025570A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Gesner, Thomas G.
;
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
;
; US-11-169-232-46
```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/11/169,232
; FILING DATE: 28-Jun-2005
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-Jan-1991
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-Jun-1990
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cserr, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)876-1170
; TELEFAX: (617)876-5851
;
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1320 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 60:
;
; US-11-169-232-60
;
Query Match      86.8%; Score 33; DB 6; Length 1320;
Best Local Similarity 85.7%; Pred. No. 1.5e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
;
QY      1 WMLSAFS 7
Db      1087 WMLSPFS 1093
;
RESULT 40
US-10-124-557-48
; Sequence 48, Application US/10124557
; Publication No. US20020137894A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Gesner, Thomas G.
;
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-Jan-1991
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-Jun-1990
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cserr, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)876-1170
; TELEFAX: (617)876-5851
;
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1320 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 60:
;
; US-11-169-232-60
```


APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Cserr, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 876-1170
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 1354 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 48:
US-10-124-557-48

Query Match 86.8%; Score 33; DB 4; Length 1354;
Best Local Similarity 85.7%; Pred. No. 1.6e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSPFS 7
Db 1121 WMLSPFS 1127

RESULT 41
US-11-169-232-48
Sequence 48, Application US/11169232
Publication No. US2006002570A1
GENERAL INFORMATION:
APPLICANT: Turner, Katherine
Clark, Stephen C.
Jacobs, Kenneth
Hewick, Rodney M.
Gesner, Thomas G.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/11/169,232
FILING DATE: 28-Jun-2005
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/10/124,557
FILING DATE: 16-Apr-2002
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 876-1170
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 1361 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 40:
US-10-124-557-40

TELEPHONE: (617) 876-1170
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 1354 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 48:
US-11-169-232-48

Query Match 86.8%; Score 33; DB 6; Length 1354;
Best Local Similarity 85.7%; Pred. No. 1.6e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSPFS 7
Db 1121 WMLSPFS 1127

RESULT 42
US-10-124-557-40
Sequence 40, Application US/10124557
Publication No. US20020137894A1
GENERAL INFORMATION:
APPLICANT: Turner, Katherine
Clark, Stephen C.
Jacobs, Kenneth
Hewick, Rodney M.
Gesner, Thomas G.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/124,557
FILING DATE: 16-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Cserr, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 876-1170
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 1361 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 40:
US-10-124-557-40

Query Match 86.8%; Score 33; DB 4; Length 1361;
Best Local Similarity 85.7%; Pred. No. 1.6e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 WMLSAFS 7
Db 1128 WMLSPFS 1134

RESULT 43
US-11-169-232-40
; Sequence 40, Application US/11169232
; Publication No. US20060025570A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Gesner, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/11/169,232
; FILING DATE: 28-Jun-2005
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-Jan-1991
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-Jun-1990
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-Dec-1989
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-Aug-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cserf, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)876-1170
; TELEFAX: (617)876-5851
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1361 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 40:
US-11-169-232-40

Query Match 86.8%; Score 33; DB 6; Length 1361;
Best Local Similarity 85.7%; Pred. No. 1.6e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 WMLSAFS 7
Db 1128 WMLSPFS 1134

RESULT 44
US-10-124-557-52
; Sequence 52, Application US/10124557
; Publication No. US20020137894A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Gesner, Thomas G.

; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-Jan-1991
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-Jun-1990
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-Dec-1989
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-Aug-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cserf, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)876-1170
; TELEFAX: (617)876-5851
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1363 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 52:
US-10-124-557-52

Query Match 86.8%; Score 33; DB 4; Length 1363;
Best Local Similarity 85.7%; Pred. No. 1.6e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 WMLSAFS 7
Db 1130 WMLSPFS 1136

RESULT 45
US-11-169-232-52
; Sequence 52, Application US/11169232
; Publication No. US20060025570A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Gesner, Thomas G.

; TITLE OF INVENTION: Megakaryocyte Stimulating Factors

```
;
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/11/169,232
; FILING DATE: 28-Jun-2005
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Geert, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 876-1170
; TELEFAX: (617) 876-5851
;
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1363 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 52:
;
; US-11-169-232-52
;
; Query Match      86.8%; Score 33; DB 6; Length 1363;
; Best Local Similarity 85.7%; Pred. No. 1.6e+03;
; Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
;
; QY      1 WMLSPFS 7
;         |||||
; Db      1130 WMLSPFS 1136
;
; RESULT 46
; US-09-802-207-30
; Sequence 30, Application US/09802207
; Publication No. US20020086824A1
; GENERAL INFORMATION:
; APPLICANT: Warman, Matthew
; APPLICANT: Carpen, John
; APPLICANT: Trent, Jeffrey
; APPLICANT: Marcelino, Jose
; TITLE OF INVENTION: Novel Methods and Reagents for the Treatment of Osteoarthritis
; FILE REFERENCE: Case-06212
; CURRENT APPLICATION NUMBER: US/09/802,207
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 09/619,175
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 60/145,328
; PRIOR FILING DATE: 1999-07-23
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 30
```

```
;
; LENGTH: 1404
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-802-207-30
;
; Query Match      86.8%; Score 33; DB 3; Length 1404;
; Best Local Similarity 85.7%; Pred. No. 1.6e+03;
; Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
;
; QY      1 WMLSPFS 7
;         |||||
; Db      1171 WMLSPFS 1177
;
; RESULT 47
; US-09-897-188-1
; Sequence 1, Application US/09897188
; Publication No. US20040072741A1
; GENERAL INFORMATION:
; APPLICANT: Jay, Gregory D.
; TITLE OF INVENTION: Tribolectin Polypeptides and Uses Thereof
; FILE REFERENCE: 21486-026 CIP2
; CURRENT APPLICATION NUMBER: US/09/897,188
; CURRENT FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: 09/298,970
; PRIOR FILING DATE: 1999-04-23
; PRIOR APPLICATION NUMBER: 09/556,246
; PRIOR FILING DATE: 2000-04-24
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1404
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-897-188-1
;
; Query Match      86.8%; Score 33; DB 3; Length 1404;
; Best Local Similarity 85.7%; Pred. No. 1.6e+03;
; Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
;
; QY      1 WMLSPFS 7
;         |||||
; Db      1171 WMLSPFS 1177
;
; RESULT 48
; US-10-124-557-2
; Sequence 2, Application US/10124557
; Publication No. US20020137894A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; APPLICANT: Clark, Stephen C.
; APPLICANT: Jacobs, Kenneth
; APPLICANT: Hewick, Rodney M.
; APPLICANT: Geener, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; CLASSIFICATION: <Unknown>
```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Cserr, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)876-1170
TELEFAX: (617)876-5851
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1404 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-124-557-2

Query Match 86.8%; Score 33; DB 4; Length 1404;
Best Local Similarity 85.7%; Pred. No. 1.6e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSPFS 7
Db 1171 WMLSPFS 1177

RESULT 49
US-10-124-557-62
Sequence 62, Application US/10124557
Publication No. US20020137894A1
GENERAL INFORMATION:
APPLICANT: Turner, Katherine
Clark, Stephen C.
Jacobs, Kenneth
Hewick, Rodney M.
Gesner, Thomas G.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/124,557
FILING DATE: 16-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Cserr, Luann

REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)876-1170
TELEFAX: (617)876-5851
INFORMATION FOR SEQ ID NO: 62:
SEQUENCE CHARACTERISTICS:
LENGTH: 1404 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 62:
US-10-124-557-62

Query Match 86.8%; Score 33; DB 4; Length 1404;
Best Local Similarity 85.7%; Pred. No. 1.6e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSPFS 7
Db 1171 WMLSPFS 1177

RESULT 50
US-10-858-595-1
Sequence 1, Application US/10858595
Publication No. US20040229804A1
GENERAL INFORMATION:
APPLICANT: Jay, Gregory D.
TITLE OF INVENTION: TRIBONECTINS
FILE REFERENCE: 21486-026
CURRENT APPLICATION NUMBER: US/10/858,595
CURRENT FILING DATE: 2004-06-01
PRIOR APPLICATION NUMBER: US/09/298,970
PRIOR FILING DATE: 1999-04-23
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastsEQ for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1404
TYPE: PRT
ORGANISM: Homo sapiens
US-10-858-595-1

Query Match 86.8%; Score 33; DB 5; Length 1404;
Best Local Similarity 85.7%; Pred. No. 1.6e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSPFS 7
Db 1171 WMLSPFS 1177

Search completed: August 29, 2006, 10:18:50
Job time : 87.6931 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2006, 09:43:08 ; Search time 5.54455 Seconds
(without alignments)
86.383 Million cell updates/sec

Title: US-10-541-343-3

Perfect score: 38
Sequence: 1 WMLSAFS 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 247503 seqs, 68422524 residues

Total number of hits satisfying chosen parameters: 247503

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 500 summaries

Database :

Published Applications AA.New:*
1: /EMC_Celerra_SIDS3/prodata/2/pubppa/US09_NEW_PUB.pep:*
2: /EMC_Celerra_SIDS3/prodata/2/pubppa/US06_NEW_PUB.pep:*
3: /EMC_Celerra_SIDS3/prodata/2/pubppa/US07_NEW_PUB.pep:*
4: /EMC_Celerra_SIDS3/prodata/2/pubppa/US08_NEW_PUB.pep:*
5: /EMC_Celerra_SIDS3/prodata/2/pubppa/PCT_NEW_PUB.pep:*
6: /EMC_Celerra_SIDS3/prodata/2/pubppa/US10_NEW_PUB.pep:*
7: /EMC_Celerra_SIDS3/prodata/2/pubppa/US11_NEW_PUB.pep:*
8: /EMC_Celerra_SIDS3/prodata/2/pubppa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	81.6	301	US-10-449-902-28641	Sequence 28641, A
2	31	81.6	302	US-10-449-902-38264	Sequence 38264, A
3	31	81.6	302	US-11-174-307B-5396	Sequence 5396, Ap
4	31	81.6	371	US-11-399-675-14	Sequence 14, Appl
5	31	81.6	371	US-11-353-390A-14	Sequence 14, Appl
6	31	81.6	371	US-11-354-310A-14	Sequence 14, Appl
7	30	78.9	504	US-10-449-902-52867	Sequence 52867, A
8	30	78.9	504	US-10-449-902-46155	Sequence 46155, A
9	29	76.3	63	US-10-449-902-53284	Sequence 53284, A
10	29	76.3	95	US-11-056-355B-79384	Sequence 79384, A
11	29	76.3	297	US-10-449-902-44647	Sequence 44647, A
12	29	76.3	308	US-10-953-349-38417	Sequence 38417, A
13	29	76.3	308	US-11-056-355B-79383	Sequence 79383, A
14	29	76.3	334	US-11-056-355B-82525	Sequence 82525, A
15	29	76.3	335	US-10-511-455-86	Sequence 26, Appl
16	29	76.3	335	US-11-395-249-4	Sequence 4, Appl1
17	29	76.3	338	US-11-056-355B-82529	Sequence 82529, A
18	29	76.3	345	US-10-953-349-33299	Sequence 33299, A
19	29	76.3	345	US-10-953-349-38416	Sequence 38416, A
20	29	76.3	351	US-11-056-355B-79382	Sequence 79382, A
21	29	76.3	354	US-10-449-902-30511	Sequence 30511, A
22	29	76.3	354	US-10-449-902-49957	Sequence 49957, A
23	29	76.3	355	US-10-953-349-32815	Sequence 32815, A
24	29	76.3	355	US-11-056-355B-11231	Sequence 11231, A
25	29	76.3	355	US-11-056-355B-17746	Sequence 17746, A

26	29	76.3	357	US-10-953-349-33298	Sequence 33298, A
27	29	76.3	357	US-10-953-349-38415	Sequence 38415, A
28	29	76.3	367	US-11-330-403-6594	Sequence 6594, Ap
29	29	76.3	368	US-10-449-902-38671	Sequence 38671, A
30	29	76.3	417	US-10-953-349-33297	Sequence 33297, A
31	29	76.3	418	US-10-953-349-32813	Sequence 32813, A
32	29	76.3	418	US-11-056-355B-11229	Sequence 11229, A
33	29	76.3	418	US-11-056-355B-17744	Sequence 17744, A
34	29	76.3	437	US-11-056-355B-82524	Sequence 82524, A
35	29	76.3	437	US-11-056-355B-82528	Sequence 82528, A
36	29	76.3	495	US-10-953-349-31757	Sequence 31757, A
37	29	76.3	495	US-11-056-355B-66755	Sequence 66755, A
38	29	76.3	530	US-10-953-349-31756	Sequence 31756, A
39	29	76.3	530	US-11-056-355B-66754	Sequence 66754, A
40	29	76.3	564	US-11-056-355B-82523	Sequence 82523, A
41	29	76.3	565	US-11-056-355B-82527	Sequence 82527, A
42	29	76.3	599	US-10-449-902-47030	Sequence 47030, A
43	29	76.3	610	US-10-953-349-31755	Sequence 31755, A
44	29	76.3	610	US-11-056-355B-66753	Sequence 66753, A
45	29	76.3	638	US-10-953-349-33831	Sequence 33831, A
46	29	76.3	665	US-10-953-349-33830	Sequence 33830, A
47	29	76.3	667	US-10-953-349-33829	Sequence 33829, A
48	28	73.7	105	US-11-056-355B-1496	Sequence 1496, Ap
49	28	73.7	124	US-11-056-355B-1495	Sequence 1495, Ap
50	28	73.7	154	US-11-056-355B-1494	Sequence 1494, Ap
51	28	73.7	214	US-11-056-355B-38256	Sequence 38256, A
52	28	73.7	214	US-11-056-355B-103593	Sequence 103593, A
53	28	73.7	214	US-11-056-355B-114832	Sequence 114832, A
54	28	73.7	219	US-11-056-355B-114831	Sequence 114831, A
55	28	73.7	250	US-10-449-902-38657	Sequence 38657, A
56	28	73.7	261	US-11-056-355B-10692	Sequence 10692, A
57	28	73.7	281	US-11-056-355B-10691	Sequence 10691, A
58	28	73.7	284	US-10-953-349-8449	Sequence 8449, Ap
59	28	73.7	291	US-10-953-349-8448	Sequence 8448, Ap
60	28	73.7	302	US-11-056-355B-36255	Sequence 36255, A
61	28	73.7	302	US-11-056-355B-103592	Sequence 103592, A
62	28	73.7	302	US-11-056-355B-114831	Sequence 114831, A
63	28	73.7	304	US-10-953-349-35221	Sequence 35221, A
64	28	73.7	304	US-11-056-355B-3247	Sequence 3247, Ap
65	28	73.7	305	US-10-953-349-8447	Sequence 8447, Ap
66	28	73.7	306	US-10-953-349-35220	Sequence 35220, A
67	28	73.7	306	US-11-056-355B-3246	Sequence 3246, Ap
68	28	73.7	322	US-11-056-355B-38254	Sequence 38254, A
69	28	73.7	322	US-11-056-355B-103591	Sequence 103591, A
70	28	73.7	322	US-11-056-355B-114830	Sequence 114830, A
71	28	73.7	348	US-10-953-349-3741	Sequence 3741, Ap
72	28	73.7	348	US-10-953-349-3740	Sequence 3740, Ap
73	28	73.7	372	US-11-399-675-12	Sequence 12, Appl
74	28	73.7	372	US-11-353-390A-12	Sequence 12, Appl
75	28	73.7	372	US-11-354-310A-12	Sequence 12, Appl
76	28	73.7	372	US-11-197-712-255	Sequence 255, App
77	28	73.7	424	US-10-953-349-3739	Sequence 3739, Ap
78	28	73.7	429	US-11-330-403-6100	Sequence 6100, Ap
79	28	73.7	439	US-11-330-403-16165	Sequence 16165, A
80	28	73.7	450	US-10-953-349-23284	Sequence 23284, A
81	28	73.7	450	US-11-056-355B-56941	Sequence 56941, A
82	28	73.7	450	US-11-330-403-1803	Sequence 1803, Ap
83	28	73.7	460	US-10-953-349-23283	Sequence 23283, A
84	28	73.7	460	US-11-056-355B-56940	Sequence 56940, A
85	28	73.7	468	US-10-449-902-46728	Sequence 46728, A
86	28	73.7	499	US-10-449-902-46724	Sequence 46724, A
87	28	73.7	508	US-10-449-902-44377	Sequence 44377, A
88	28	73.7	527	US-10-449-902-46271	Sequence 46271, A
89	28	73.7	537	US-10-539-228-787	Sequence 787, App
90	28	73.7	605	US-10-449-902-45413	Sequence 45413, A
91	28	73.7	658	US-10-449-902-38278	Sequence 38278, A
92	28	73.7	675	US-10-449-902-42270	Sequence 42270, A
93	28	73.7	805	US-10-449-902-53596	Sequence 53596, A
94	28	73.7	830	US-10-535-928-35	Sequence 35, Appl
95	27	71.1	1099	US-10-196-749-266	Sequence 266, App
96	27	71.1	151	US-11-056-355B-72635	Sequence 72635, A
97	27	71.1	162	US-11-197-712-391	Sequence 391, App
98	27	71.1	208	US-10-953-349-17085	Sequence 17085, A
99	27	71.1	210	US-10-953-349-17084	Sequence 17084, A

99	27	71.1	223	6	US-10-953-349-17083	Sequence 17083, A	172	26	68.4	133	7	US-11-056-355B-14590	Sequence 14590, A
100	27	71.1	242	7	US-11-056-355B-90138	Sequence 90138, A	173	26	68.4	155	6	US-10-953-349-10769	Sequence 10769, A
101	27	71.1	242	7	US-11-056-355B-93894	Sequence 93894, A	174	26	68.4	157	6	US-10-953-349-10768	Sequence 10768, A
102	27	71.1	253	7	US-11-330-403-7966	Sequence 7966, Ap	175	26	68.4	160	6	US-10-953-349-10767	Sequence 10767, A
103	27	71.1	256	7	US-11-056-355B-90137	Sequence 90137, A	176	26	68.4	155	7	US-11-056-355B-56901	Sequence 56901, A
104	27	71.1	256	7	US-11-056-355B-93893	Sequence 93893, A	177	26	68.4	168	7	US-11-293-697-2902	Sequence 2902, Ap
105	27	71.1	308	6	US-10-449-902-32114	Sequence 32114, A	178	26	68.4	163	6	US-10-953-349-33264	Sequence 33264, A
106	27	71.1	311	7	US-11-174-307B-34422	Sequence 34422, Ap	179	26	68.4	163	6	US-10-953-349-35669	Sequence 35669, A
107	27	71.1	315	6	US-10-953-349-2044	Sequence 2044, Ap	180	26	68.4	133	7	US-11-056-355B-14289	Sequence 14289, A
108	27	71.1	321	6	US-10-953-349-2043	Sequence 2043, Ap	181	26	68.4	163	7	US-11-056-355B-14589	Sequence 14589, A
109	27	71.1	321	6	US-10-449-902-43307	Sequence 43307, A	182	26	68.4	166	6	US-10-953-349-11510	Sequence 11510, A
110	27	71.1	326	6	US-10-953-349-2042	Sequence 2042, Ap	183	26	68.4	205	6	US-10-507-908-9	Sequence 9, Appl1
111	27	71.1	343	6	US-11-056-355B-7695	Sequence 7695, Ap	184	26	68.4	206	6	US-10-953-349-15493	Sequence 15493, A
112	27	71.1	345	7	US-11-056-355B-7694	Sequence 7694, Ap	185	26	68.4	209	6	US-10-953-349-11509	Sequence 11509, A
113	27	71.1	349	6	US-10-449-902-51199	Sequence 51199, A	186	26	68.4	218	7	US-11-056-355B-107185	Sequence 107185, A
114	27	71.1	355	7	US-11-330-403-14500	Sequence 14500, A	187	26	68.4	218	7	US-11-056-355B-118424	Sequence 118424, A
115	27	71.1	360	7	US-11-056-355B-53811	Sequence 53811, A	188	26	68.4	224	6	US-10-953-349-15492	Sequence 15492, A
116	27	71.1	367	7	US-11-199-489A-16	Sequence 16, Appl	189	26	68.4	220	7	US-11-056-355B-56900	Sequence 56900, A
117	27	71.1	380	7	US-11-056-355B-69940	Sequence 69940, A	190	26	68.4	232	7	US-11-056-355B-76032	Sequence 76032, A
118	27	71.1	396	7	US-11-056-355B-57428	Sequence 57428, A	191	26	68.4	245	6	US-10-953-349-33263	Sequence 33263, A
119	27	71.1	396	6	US-10-953-349-10976	Sequence 10976, A	192	26	68.4	245	6	US-10-953-349-35668	Sequence 35668, A
120	27	71.1	407	7	US-11-056-355B-69939	Sequence 69939, A	193	26	68.4	245	7	US-11-056-355B-14288	Sequence 14288, A
121	27	71.1	409	7	US-11-056-355B-7693	Sequence 7693, Ap	194	26	68.4	245	7	US-11-056-355B-14588	Sequence 14588, A
122	27	71.1	417	7	US-11-056-355B-53810	Sequence 53810, A	195	26	68.4	249	7	US-11-056-355B-1429	Sequence 1429, Ap
123	27	71.1	423	6	US-10-471-571A-5438	Sequence 5438, Ap	196	26	68.4	256	7	US-11-056-355B-30921	Sequence 30921, Ap
124	27	71.1	447	7	US-11-056-355B-57427	Sequence 57427, A	197	26	68.4	256	7	US-11-056-355B-34511	Sequence 34511, A
125	27	71.1	449	7	US-11-316-535A-42	Sequence 42, Appl	198	26	68.4	256	7	US-11-056-355B-101446	Sequence 101446, A
126	27	71.1	457	7	US-11-056-355B-57426	Sequence 57426, A	199	26	68.4	256	7	US-11-056-355B-112685	Sequence 112685, A
127	27	71.1	474	7	US-11-241-596-46	Sequence 46, Appl	200	26	68.4	258	7	US-11-056-355B-10920	Sequence 10920, A
128	27	71.1	474	7	US-11-056-355B-41124	Sequence 41124, A	201	26	68.4	258	7	US-11-056-355B-34510	Sequence 34510, A
129	27	71.1	474	7	US-11-056-355B-47702	Sequence 47702, A	202	26	68.4	258	7	US-11-056-355B-101445	Sequence 101445, A
130	27	71.1	485	6	US-10-449-902-48086	Sequence 48086, A	203	26	68.4	258	7	US-11-056-355B-112684	Sequence 112684, A
131	27	71.1	486	6	US-10-449-902-50373	Sequence 50373, A	204	26	68.4	259	7	US-11-056-355B-73089	Sequence 73089, A
132	27	71.1	488	7	US-11-056-355B-69938	Sequence 69938, A	205	26	68.4	260	7	US-11-056-355B-76031	Sequence 76031, A
133	27	71.1	493	6	US-10-953-349-10975	Sequence 10975, A	206	26	68.4	261	7	US-11-056-355B-30919	Sequence 30919, A
134	27	71.1	502	7	US-11-296-657-2	Sequence 2, Appl1	207	26	68.4	261	7	US-11-056-355B-34509	Sequence 34509, A
135	27	71.1	508	7	US-11-296-657-3	Sequence 3, Appl1	208	26	68.4	261	7	US-11-056-355B-73088	Sequence 73088, A
136	27	71.1	516	6	US-10-449-902-40874	Sequence 40874, A	209	26	68.4	261	7	US-11-056-355B-101444	Sequence 101444, A
137	27	71.1	516	7	US-11-296-657-5	Sequence 5, Appl1	210	26	68.4	261	7	US-11-056-355B-112683	Sequence 112683, A
138	27	71.1	520	7	US-11-174-307B-5508	Sequence 5508, Ap	211	26	68.4	265	7	US-11-056-355B-107184	Sequence 107184, A
139	27	71.1	520	7	US-11-296-657-4	Sequence 4, Appl1	212	26	68.4	265	7	US-11-056-355B-118423	Sequence 118423, A
140	27	71.1	541	7	US-11-241-586-45	Sequence 45, Appl1	213	26	68.4	272	7	US-11-056-355B-107183	Sequence 107183, A
141	27	71.1	541	7	US-11-056-355B-41123	Sequence 41123, A	214	26	68.4	272	7	US-11-056-355B-118422	Sequence 118422, A
142	27	71.1	541	7	US-11-056-355B-47701	Sequence 47701, A	215	26	68.4	283	7	US-11-056-355B-1428	Sequence 1428, Ap
143	27	71.1	543	6	US-10-449-902-43794	Sequence 43794, A	216	26	68.4	284	6	US-10-449-902-52300	Sequence 52300, A
144	27	71.1	583	7	US-11-241-596-44	Sequence 44, Appl1	217	26	68.4	304	6	US-10-449-902-56374	Sequence 56374, A
145	27	71.1	583	7	US-11-056-355B-41122	Sequence 41122, A	218	26	68.4	308	7	US-11-331-240-31	Sequence 31, Appl1
146	27	71.1	583	7	US-11-056-355B-47700	Sequence 47700, A	219	26	68.4	310	7	US-11-056-355B-76630	Sequence 76630, A
147	27	71.1	647	7	US-11-250-089-50	Sequence 50, Appl1	220	26	68.4	319	6	US-10-449-902-42809	Sequence 42809, A
148	27	71.1	712	6	US-11-250-089-52	Sequence 52, Appl1	221	26	68.4	331	6	US-10-449-902-53432	Sequence 53432, A
149	27	71.1	715	6	US-10-199-229-15	Sequence 15, Appl1	222	26	68.4	331	6	US-10-449-902-55653	Sequence 55653, A
150	27	71.1	830	6	US-10-449-902-41565	Sequence 41565, A	223	26	68.4	333	6	US-10-953-349-15724	Sequence 15724, A
151	27	71.1	936	6	US-10-449-902-52526	Sequence 52526, A	224	26	68.4	333	6	US-11-056-355B-55257	Sequence 55257, A
152	27	71.1	975	6	US-10-449-902-37603	Sequence 37603, A	225	26	68.4	334	6	US-10-953-349-15723	Sequence 15723, A
153	27	71.1	1181	7	US-11-241-596-355	Sequence 255, Appl	226	26	68.4	334	7	US-11-056-355B-55356	Sequence 55356, A
154	27	71.1	1181	7	US-11-241-596-256	Sequence 256, Appl	227	26	68.4	342	7	US-11-056-355B-1427	Sequence 1427, Ap
155	27	71.1	1181	7	US-11-241-596-257	Sequence 257, Appl	228	26	68.4	345	7	US-11-330-403-4631	Sequence 4631, Ap
156	27	71.1	1181	7	US-11-241-596-258	Sequence 258, Appl	229	26	68.4	346	7	US-11-376-694-1	Sequence 1, Appl1
157	27	71.1	1403	6	US-10-505-928-471	Sequence 471, Appl	230	26	68.4	349	7	US-11-404-939-503	Sequence 503, Appl
158	27	71.1	1601	6	US-10-539-228-225	Sequence 225, Appl	231	26	68.4	361	7	US-11-056-355B-19584	Sequence 19584, A
159	27	71.1	1625	7	US-11-368-804-2	Sequence 2, Appl1	232	26	68.4	393	6	US-10-449-902-28875	Sequence 28875, A
160	26	68.4	51	7	US-11-056-355B-50950	Sequence 50950, A	233	26	68.4	393	6	US-10-449-902-43261	Sequence 43261, A
161	26	68.4	52	7	US-11-115-024-26	Sequence 26, Appl1	234	26	68.4	396	6	US-10-449-902-32606	Sequence 32606, A
162	26	68.4	79	6	US-11-115-024-290	Sequence 290, Appl	235	26	68.4	398	7	US-11-330-403-9534	Sequence 9534, Ap
163	26	68.4	79	6	US-10-953-349-29615	Sequence 29615, A	236	26	68.4	403	7	US-11-174-307B-2752	Sequence 2752, Appl
164	26	68.4	82	6	US-10-449-902-53647	Sequence 53647, A	237	26	68.4	430	7	US-11-155-866-75	Sequence 75, Appl1
165	26	68.4	86	6	US-10-449-902-56450	Sequence 56450, A	238	26	68.4	433	7	US-11-330-403-10896	Sequence 10896, A
166	26	68.4	106	6	US-10-449-902-32052	Sequence 32052, A	239	26	68.4	433	7	US-11-330-403-15322	Sequence 15322, A
167	26	68.4	111	6	US-10-953-349-29614	Sequence 29614, A	240	26	68.4	438	6	US-10-449-902-47924	Sequence 47924, A
168	26	68.4	127	6	US-11-056-355B-56902	Sequence 56902, A	241	26	68.4	439	6	US-10-953-349-4428	Sequence 4428, Ap
169	26	68.4	133	6	US-10-953-349-33265	Sequence 33265, A	242	26	68.4	439	7	US-11-056-355B-41492	Sequence 41492, A
170	26	68.4	133	6	US-10-953-349-35670	Sequence 35670, A	243	26	68.4	439	7	US-11-056-355B-96486	Sequence 96486, A
171	26	68.4	133	7	US-11-056-355B-14290	Sequence 14290, A	244	26	68.4	443	6	US-10-511-937-2464	Sequence 2464, Ap

245	26	68.4	443	7	US-11-289-102-300	Sequence 300, App	318	25	65.8	166	6	US-10-449-902-30822	Sequence 30822, A
246	26	68.4	443	7	US-11-289-102-354	Sequence 354, App	319	25	65.8	179	7	US-11-056-355B-58360	Sequence 58360, A
247	26	68.4	443	7	US-11-289-102-379	Sequence 379, App	320	25	65.8	191	6	US-10-953-349-4195	Sequence 4195, App
248	26	68.4	444	6	US-10-953-349-4427	Sequence 4427, App	321	25	65.8	191	7	US-11-056-355B-44513	Sequence 44513, A
249	26	68.4	444	7	US-11-056-355B-34829	Sequence 34829, A	322	25	65.8	191	7	US-11-056-355B-98975	Sequence 98975, A
250	26	68.4	444	7	US-11-056-355B-41491	Sequence 41491, A	323	25	65.8	191	7	US-11-056-355B-110214	Sequence 110214, A
251	26	68.4	444	7	US-11-056-355B-96485	Sequence 96485, A	324	25	65.8	206	7	US-11-056-355B-58359	Sequence 58359, A
252	26	68.4	448	7	US-11-024-544-136	Sequence 136, App	325	25	65.8	211	7	US-11-056-355B-30601	Sequence 30601, A
253	26	68.4	458	7	US-11-190-750-119	Sequence 119, App	326	25	65.8	211	7	US-11-056-355B-34191	Sequence 34191, A
254	26	68.4	458	7	US-11-264-784-5	Sequence 5, App1	327	25	65.8	220	6	US-10-953-349-1110	Sequence 1110, App
255	26	68.4	458	7	US-11-264-737-5	Sequence 5, App1	328	25	65.8	220	6	US-10-449-902-39665	Sequence 39665, A
256	26	68.4	458	7	US-11-265-761-5	Sequence 5, App1	329	25	65.8	221	7	US-11-056-355B-30600	Sequence 30600, A
257	26	68.4	461	6	US-10-449-902-40732	Sequence 40732, A	330	25	65.8	221	7	US-11-056-355B-34190	Sequence 34190, A
258	26	68.4	434	7	US-11-056-355B-78762	Sequence 78762, A	331	25	65.8	225	6	US-10-196-749-358	Sequence 358, App
259	26	68.4	504	7	US-11-056-355B-54417	Sequence 54417, A	332	25	65.8	225	7	US-10-196-712-393	Sequence 393, App
260	26	68.4	509	7	US-11-056-355B-73428	Sequence 73428, A	333	25	65.8	225	7	US-11-101-316-120	Sequence 120, App
261	26	68.4	517	6	US-10-449-902-40592	Sequence 40592, A	334	25	65.8	225	7	US-11-197-712-405	Sequence 405, App
262	26	68.4	524	7	US-11-056-355B-73427	Sequence 73427, A	335	25	65.8	225	7	US-11-376-673-120	Sequence 120, App
263	26	68.4	527	7	US-11-056-355B-71228	Sequence 71228, A	336	25	65.8	226	6	US-10-953-349-4194	Sequence 4194, App
264	26	68.4	532	7	US-11-056-355B-54416	Sequence 54416, A	337	25	65.8	226	6	US-11-056-355B-41512	Sequence 41512, A
265	26	68.4	532	7	US-11-056-355B-71227	Sequence 71227, A	338	25	65.8	226	7	US-11-056-355B-98974	Sequence 98974, A
266	26	68.4	533	7	US-11-056-355B-83566	Sequence 83566, A	339	25	65.8	226	7	US-11-056-355B-110213	Sequence 110213, A
267	26	68.4	554	7	US-11-056-355B-96484	Sequence 96484, A	340	25	65.8	250	6	US-10-953-349-1109	Sequence 1109, App
268	26	68.4	559	6	US-10-953-349-4426	Sequence 4426, App	341	25	65.8	250	6	US-10-449-902-50550	Sequence 50550, A
269	26	68.4	559	7	US-11-056-355B-34828	Sequence 34828, A	342	25	65.8	250	6	US-10-449-902-17884	Sequence 17884, A
270	26	68.4	559	7	US-11-056-355B-41490	Sequence 41490, A	343	25	65.8	250	7	US-11-056-355B-27674	Sequence 27674, A
271	26	68.4	572	6	US-10-449-902-38041	Sequence 38041, A	344	25	65.8	270	7	US-11-056-355B-30599	Sequence 30599, A
272	26	68.4	572	6	US-10-449-902-53023	Sequence 53023, A	345	25	65.8	270	7	US-11-056-355B-31264	Sequence 31264, A
273	26	68.4	572	6	US-11-292-951-8	Sequence 8, App1	346	25	65.8	270	7	US-11-056-355B-34189	Sequence 34189, A
274	26	68.4	577	7	US-11-056-355B-83565	Sequence 83565, A	347	25	65.8	270	7	US-11-056-355B-74297	Sequence 74297, A
275	26	68.4	579	7	US-11-292-951-6	Sequence 6, App1	348	25	65.8	273	6	US-10-449-902-29449	Sequence 29499, A
276	26	68.4	579	7	US-11-292-951-4	Sequence 4, App1	349	25	65.8	275	6	US-10-540-898-639	Sequence 639, App
277	26	68.4	579	7	US-11-056-355B-34827	Sequence 34827, A	350	25	65.8	277	7	US-11-056-355B-1636	Sequence 1636, App
278	26	68.4	579	7	US-11-330-403-9989	Sequence 9, App1	351	25	65.8	280	7	US-11-056-355B-58358	Sequence 58358, A
279	26	68.4	580	7	US-11-292-951-9	Sequence 9, App1	352	25	65.8	289	6	US-10-449-902-38640	Sequence 38640, A
280	26	68.4	580	7	US-11-296-657-46	Sequence 46, App1	353	25	65.8	293	7	US-11-330-403-3736	Sequence 3736, App
281	26	68.4	580	7	US-11-330-403-17079	Sequence 17079, A	354	25	65.8	296	6	US-10-953-349-12428	Sequence 12428, A
282	26	68.4	585	6	US-10-449-902-45375	Sequence 45375, A	355	25	65.8	298	6	US-11-330-403-8484	Sequence 8484, App
283	26	68.4	585	6	US-10-449-902-55150	Sequence 55150, A	356	25	65.8	299	6	US-10-449-902-29085	Sequence 29085, A
284	26	68.4	591	7	US-11-045-540-1	Sequence 1, App1	357	25	65.8	308	6	US-10-953-349-27488	Sequence 27488, A
285	26	68.4	591	7	US-11-056-355B-54415	Sequence 54415, A	358	25	65.8	310	7	US-11-056-355B-64504	Sequence 64504, A
286	26	68.4	602	7	US-11-056-355B-73426	Sequence 73426, A	359	25	65.8	310	7	US-11-056-355B-1635	Sequence 1635, App
287	26	68.4	604	7	US-11-056-355B-83564	Sequence 83564, A	360	25	65.8	314	6	US-11-056-355B-1634	Sequence 1634, App
288	26	68.4	615	7	US-11-330-403-6742	Sequence 6742, App	361	25	65.8	315	6	US-10-953-349-12427	Sequence 12427, A
289	26	68.4	620	7	US-11-056-355B-71226	Sequence 71226, A	362	25	65.8	316	7	US-11-330-403-13413	Sequence 13413, A
290	26	68.4	621	7	US-11-199-489A-120	Sequence 120, App	363	25	65.8	317	6	US-10-953-349-27487	Sequence 27487, A
291	26	68.4	625	6	US-10-449-902-41455	Sequence 41455, A	364	25	65.8	317	6	US-11-056-355B-64503	Sequence 64503, A
292	26	68.4	672	6	US-10-370-959-154	Sequence 154, App	365	25	65.8	322	7	US-11-387-009-2	Sequence 2, App1
293	26	68.4	698	6	US-10-449-902-44559	Sequence 44559, A	366	25	65.8	322	7	US-11-293-697-4057	Sequence 4057, App
294	26	68.4	788	6	US-10-539-928-663	Sequence 663, App	367	25	65.8	331	6	US-10-449-902-34863	Sequence 34863, A
295	26	68.4	801	6	US-10-539-928-666	Sequence 666, App	368	25	65.8	334	7	US-11-056-355B-85829	Sequence 85829, A
296	26	68.4	1194	7	US-11-045-540-2	Sequence 2, App1	369	25	65.8	337	7	US-11-056-355B-73681	Sequence 73681, A
297	26	68.4	1205	7	US-11-045-540-3	Sequence 3, App1	370	25	65.8	342	7	US-11-056-355B-73680	Sequence 73680, A
298	26	68.4	1237	7	US-11-045-540-4	Sequence 4, App1	371	25	65.8	342	7	US-10-471-571A-3982	Sequence 3982, App
299	26	68.4	1248	7	US-11-045-540-5	Sequence 5, App1	372	25	65.8	350	6	US-10-449-902-45868	Sequence 45868, A
300	26	68.4	1346	6	US-10-449-902-55021	Sequence 55021, A	373	25	65.8	350	7	US-11-330-403-17229	Sequence 17229, A
301	26	68.4	1397	6	US-10-506-630A-36	Sequence 36, App1	374	25	65.8	355	7	US-11-056-355B-47544	Sequence 47544, A
302	26	68.4	2627	6	US-10-547-530-5	Sequence 5, App1	375	25	65.8	355	7	US-11-056-355B-47602	Sequence 47602, A
303	26	68.4	10625	6	US-11-330-403-1837	Sequence 1837, App	376	25	65.8	356	6	US-10-953-349-27486	Sequence 27486, A
304	26	68.4	22152	6	US-10-544-944-1	Sequence 1, App1	377	25	65.8	356	6	US-11-056-355B-64502	Sequence 64502, A
305	26	65.8	40	6	US-11-197-712-407	Sequence 407, App	378	25	65.8	358	6	US-10-953-349-12426	Sequence 12426, A
306	26	65.8	80	6	US-10-953-349-29217	Sequence 29217, A	379	25	65.8	358	7	US-11-056-355B-1106341	Sequence 1106341, A
307	26	65.8	104	6	US-10-953-349-29216	Sequence 29216, A	380	25	65.8	358	7	US-11-056-355B-117580	Sequence 117580, A
308	26	65.8	104	6	US-10-449-902-52060	Sequence 52060, A	381	25	65.8	360	7	US-11-056-355B-73679	Sequence 73679, A
309	26	65.8	105	6	US-11-293-697-4709	Sequence 4709, App	382	25	65.8	361	6	US-10-449-902-53487	Sequence 53487, A
310	26	65.8	125	6	US-10-449-902-38842	Sequence 38842, A	383	25	65.8	363	7	US-11-056-355B-23707	Sequence 23707, A
311	26	65.8	126	7	US-11-293-697-2502	Sequence 2502, App	384	25	65.8	363	7	US-11-056-355B-106340	Sequence 106340, A
312	26	65.8	133	6	US-10-449-902-29280	Sequence 29280, App	385	25	65.8	363	7	US-11-056-355B-117579	Sequence 117579, A
313	26	65.8	140	6	US-10-953-349-4196	Sequence 4196, App	386	25	65.8	365	6	US-10-953-349-1108	Sequence 1108, App
314	26	65.8	140	7	US-11-056-355B-41514	Sequence 41514, A	387	25	65.8	366	6	US-10-449-902-46434	Sequence 46434, App
315	26	65.8	140	7	US-11-056-355B-98976	Sequence 98976, A	388	25	65.8	370	7	US-11-056-355B-79818	Sequence 79818, A
316	26	65.8	140	7	US-11-056-355B-110215	Sequence 110215, A	389	25	65.8	376	6	US-10-449-902-49313	Sequence 49313, A
317	26	65.8	160	6	US-10-471-571A-4622	Sequence 4622, App	390	25	65.8	378	7	US-11-330-403-947	Sequence 947, App

391	25	65	8	381	6	US-10-853-349-21199	Sequence 21199, A	444	25	65	8	543	7	US-11-265-761-92	Sequence 92, Appl
392	25	65	8	381	7	US-11-056-355B-9100	Sequence 9100, Ap	445	25	65	8	549	7	US-11-274-171A-2	Sequence 2, Appl
393	25	65	8	381	7	US-11-056-355B-23706	Sequence 23706, A	446	25	65	8	552	7	US-11-330-403-17433	Sequence 17433, A
394	25	65	8	381	7	US-11-056-355B-116339	Sequence 106339, A	447	25	65	8	554	6	US-10-449-902-56063	Sequence 56063, A
395	25	65	8	381	7	US-11-056-355B-117578	Sequence 117578, A	448	25	65	8	555	7	US-11-241-596-96	Sequence 96, Appl
396	25	65	8	386	7	US-11-056-355B-63389	Sequence 63389, A	449	25	65	8	555	7	US-11-226-554-104	Sequence 104, Appl
397	25	65	8	393	7	US-11-056-355B-85828	Sequence 85828, A	450	25	65	8	555	7	US-11-226-554-105	Sequence 105, Appl
398	25	65	8	395	7	US-11-056-355B-39106	Sequence 39106, A	451	25	65	8	555	7	US-11-248-718-104	Sequence 104, Appl
399	25	65	8	396	6	US-10-449-902-50121	Sequence 50121, A	452	25	65	8	555	7	US-11-248-718-104	Sequence 105, App
400	25	65	8	401	7	US-11-056-355B-63388	Sequence 63388, A	453	25	65	8	569	7	US-11-056-355B-90206	Sequence 90206, A
401	25	65	8	407	7	US-11-056-355B-57139	Sequence 57139, A	454	25	65	8	569	7	US-11-056-355B-93962	Sequence 93962, A
402	25	65	8	409	7	US-11-056-355B-23705	Sequence 23705, A	455	25	65	8	570	7	US-11-317-789A-391	Sequence 391, App
403	25	65	8	417	7	US-11-056-355B-82642	Sequence 82642, A	456	25	65	8	574	6	US-10-449-902-81029	Sequence 41029, A
404	25	65	8	419	7	US-11-330-403-14803	Sequence 14903, A	457	25	65	8	582	7	US-11-056-355B-80484	Sequence 80484, A
405	25	65	8	423	6	US-10-953-349-2270	Sequence 22707, Ap	458	25	65	8	586	7	US-11-056-355B-80479	Sequence 80479, A
406	25	65	8	423	7	US-11-056-355B-40032	Sequence 40032, A	459	25	65	8	592	6	US-10-449-902-55047	Sequence 55047, A
407	25	65	8	425	6	US-10-449-902-29640	Sequence 40906, A	460	25	65	8	597	7	US-11-245-628-53	Sequence 43, Appl
408	25	65	8	426	6	US-10-449-902-53319	Sequence 53319, A	461	25	65	8	598	7	US-11-241-596-95	Sequence 95, Appl
409	25	65	8	429	7	US-11-056-355B-79817	Sequence 79817, A	462	25	65	8	598	7	US-11-174-307B-3672	Sequence 3672, Ap
410	25	65	8	430	6	US-10-449-902-37851	Sequence 79817, A	463	25	65	8	607	6	US-10-449-902-45927	Sequence 45927, A
411	25	65	8	430	6	US-10-449-902-37851	Sequence 79851, A	464	25	65	8	607	7	US-11-056-355B-39104	Sequence 39104, A
412	25	65	8	430	7	US-11-056-355B-9099	Sequence 9099, Ap	465	25	65	8	607	7	US-11-056-355B-80483	Sequence 80483, A
413	25	65	8	431	7	US-11-330-403-1168	Sequence 1168, Ap	466	25	65	8	610	6	US-10-449-902-33952	Sequence 33952, A
414	25	65	8	432	7	US-11-330-403-1382	Sequence 1382, Ap	467	25	65	8	611	6	US-10-449-902-82603	Sequence 42603, A
415	25	65	8	433	7	US-11-330-403-6219	Sequence 6219, Ap	468	25	65	8	619	7	US-11-312-958-46	Sequence 46, Appl
416	25	65	8	436	7	US-11-330-403-31295	Sequence 31295, A	469							

Db 192 WMLTFS 198

RESULT 2

US-10-449-902-38264
; Sequence 38264, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agricultural Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-20369
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 38264
; LENGTH: 302
; TYPE: PRT
; ORGANISM: *Oryza sativa*
US-10-449-902-38264

Query Match 81.6%; Score 31; DB 6; Length 302;
Best Local Similarity 71.4%; Pred. No. 62;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLTFS 7

Db 193 WMLTFS 199

RESULT 3

US-11-174-307B-5396
; Sequence 5396, Application US/11174307B
; Publication No. US20060143729A1
; GENERAL INFORMATION:
; APPLICANT: BROVER, Vyacheslav
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES AND POLYPEPTIDES ENCODED THEREBY
; FILE REFERENCE: 2750-1601PUS2
; CURRENT APPLICATION NUMBER: US/11/174,307B
; CURRENT FILING DATE: 2005-06-30
; PRIOR APPLICATION NUMBER: 60/583,671
; PRIOR FILING DATE: 2004-06-30
; PRIOR APPLICATION NUMBER: 60/583,781
; PRIOR FILING DATE: 2004-06-30
; PRIOR APPLICATION NUMBER: 60/583,651
; PRIOR FILING DATE: 2004-06-30
; NUMBER OF SEQ ID NOS: 5544
; SEQ ID NO 5396
; LENGTH: 302
; TYPE: PRT
; ORGANISM: *Zea mays*
US-11-174-307B-5396

Query Match 81.6%; Score 31; DB 7; Length 302;
Best Local Similarity 71.4%; Pred. No. 62;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLTFS 7

Db 193 WMLTFS 199

RESULT 4
US-11-399-675-14

; Sequence 14, Application US/11399675
; Publication No. US20060179513A1
; GENERAL INFORMATION:
; APPLICANT: Sticklen, Masomeh B
; APPLICANT: Magbool, Shahina B
; APPLICANT: Dale, Bruce E
; TITLE OF INVENTION: TRANSGENIC PLANTS CONTAINING LIGNINASE AND CELLULASE WHICH
; FILE REFERENCE: MSU 4.1-814
; CURRENT APPLICATION NUMBER: US/11/399,675
; CURRENT FILING DATE: 2006-04-06
; PRIOR APPLICATION NUMBER: US 60/242,408
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 09/981,900
; PRIOR FILING DATE: 2001-10-18
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 14
; LENGTH: 371
; TYPE: PRT
; ORGANISM: *Phanerochaete chrysosporium*
US-11-399-675-14

Query Match 81.6%; Score 31; DB 7; Length 371;
Best Local Similarity 85.7%; Pred. No. 75;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLTFS 7

Db 197 WMLTFS 203

RESULT 5

US-11-353-390A-14
; Sequence 14, Application US/11353390A
; Publication No. US20060185036A1
; GENERAL INFORMATION:
; APPLICANT: Sticklen, Masomeh B
; APPLICANT: Magbool, Shahina B
; APPLICANT: Dale, Bruce E
; TITLE OF INVENTION: TRANSGENIC PLANTS CONTAINING LIGNINASE AND CELLULASE WHICH
; FILE REFERENCE: MSU 4.1-806
; CURRENT APPLICATION NUMBER: US/11/353,390A
; CURRENT FILING DATE: 2006-02-14
; PRIOR APPLICATION NUMBER: US 60/242,408
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 09/981,900
; PRIOR FILING DATE: 2001-10-18
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 14
; LENGTH: 371
; TYPE: PRT
; ORGANISM: *Phanerochaete chrysosporium*
US-11-353-390A-14

Query Match 81.6%; Score 31; DB 7; Length 371;
Best Local Similarity 85.7%; Pred. No. 75;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLTFS 7

Db 197 WMLTFS 203

RESULT 6

US-11-354-310A-14
; Sequence 14, Application US/11354310A
; Publication No. US20060185037A1
; GENERAL INFORMATION:
; APPLICANT: Sticklen, Masomeh B
; APPLICANT: Magbool, Shahina B

```
; APPLICANT: Dale, Bruce E
; TITLE OF INVENTION: TRANSGENIC PLANTS CONTAINING LIGNINASE AND CELLULASE WHICH
; FILE REFERENCE: MEU 4.1-805
; CURRENT APPLICATION NUMBER: US/11/354.310A
; PRIOR FILING DATE: 2006-02-14/354.310A
; PRIOR APPLICATION NUMBER: US 60/242,408
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 09/981,900
; PRIOR FILING DATE: 2001-10-18
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 14
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Phanerochaete chrysosporium
US-11-354-310A-14

Query Match      81.6%; Score 31; DB 7; Length 371;
Best Local Similarity 85.7%; Pred. No. 75;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 WMLSAFS 7
Db      197 WMLSAHS 203
```

```
RESULT 7
US-10-449-902-52867
; Sequence 52867, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449.902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 52867
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-52867

Query Match      78.9%; Score 30; DB 6; Length 504;
Best Local Similarity 71.4%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 WMLSAFS 7
Db      213 WMLRAFT 219
```

```
RESULT 8
US-10-449-902-46155
; Sequence 46155, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449.902
```

```
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 46155
; LENGTH: 1128
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-46155

Query Match      78.9%; Score 30; DB 6; Length 1128;
Best Local Similarity 71.4%; Pred. No. 3.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 WMLSAFS 7
Db      451 WMLSLFN 457
```

```
RESULT 9
US-10-449-902-53284
; Sequence 53284, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449.902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 53284
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-53284

Query Match      76.3%; Score 29; DB 6; Length 63;
Best Local Similarity 57.1%; Pred. No. 34;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 WMLSAFS 7
Db      22 WMLVAFA 28
```

```
RESULT 10
US-10-449-902-44647
; Sequence 44647, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449.902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
```

NUMBER OF SEQ ID NOS: 56791
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 44647
LENGTH: 95
TYPE: PRT
ORGANISM: Oryza sativa
US-10-449-902-44647

Query Match 76.3%; Score 29; DB 6; Length 95;
Best Local Similarity 57.1%; Pred. No. 51;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
DB 54 WMLAAFA 60

RESULT 11
US-11-056-355B-79384
Sequence 79384, Application US/11056355B
Publication No. US20060150283A1
GENERAL INFORMATION:
APPLICANT: Brover, Vyacheslav
APPLICANT: Alexandrov, Nikolai
TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
FILE REFERENCE: 2750-1590PUS2
CURRENT APPLICATION NUMBER: US/11/056,355B
CURRENT FILING DATE: 2005-02-14
PRIOR APPLICATION NUMBER: 60/544,190
PRIOR FILING DATE: 2004-02-13
NUMBER OF SEQ ID NOS: 119966
SEQ ID NO 79384
LENGTH: 297
TYPE: prt
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: peptide
LOCATION: (1)..(297)
OTHER INFORMATION: Ceres Seq. ID no. 12651075
US-11-056-355B-79384

Query Match 76.3%; Score 29; DB 7; Length 297;
Best Local Similarity 71.4%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
DB 219 WMLIAFS 225

RESULT 12
US-10-953-38417
Sequence 38417, Application US/10953349
Publication No. US20060107345A1
GENERAL INFORMATION:
APPLICANT: ALEXANDROV, Nikolai et al.
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
FILE REFERENCE: 2750-1579PUS2
CURRENT APPLICATION NUMBER: US/10/953,349
CURRENT FILING DATE: 2004-09-30
NUMBER OF SEQ ID NOS: 40252
SOFTWARE: PatentIn version 3.3
SEQ ID NO 38417
LENGTH: 308
TYPE: PRT
ORGANISM: Zea mays subsp. mays
US-10-953-349-38417

Query Match 76.3%; Score 29; DB 6; Length 308;
Best Local Similarity 71.4%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
DB 233 WMLIAFS 239

RESULT 13
US-11-056-355B-79383
Sequence 79383, Application US/11056355B
Publication No. US20060150283A1
GENERAL INFORMATION:
APPLICANT: Brover, Vyacheslav
APPLICANT: Alexandrov, Nikolai
TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
FILE REFERENCE: 2750-1590PUS2
CURRENT APPLICATION NUMBER: US/11/056,355B
CURRENT FILING DATE: 2005-02-14
PRIOR APPLICATION NUMBER: 60/544,190
PRIOR FILING DATE: 2004-02-13
NUMBER OF SEQ ID NOS: 119966
SEQ ID NO 79383
LENGTH: 308
TYPE: prt
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: peptide
LOCATION: (1)..(308)
OTHER INFORMATION: Ceres Seq. ID no. 12651074
US-11-056-355B-79383

Query Match 76.3%; Score 29; DB 7; Length 308;
Best Local Similarity 71.4%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
DB 230 WMLIAFS 236

RESULT 14
US-11-056-355B-82525
Sequence 82525, Application US/11056355B
Publication No. US20060150283A1
GENERAL INFORMATION:
APPLICANT: Brover, Vyacheslav
APPLICANT: Alexandrov, Nikolai
TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
FILE REFERENCE: 2750-1590PUS2
CURRENT APPLICATION NUMBER: US/11/056,355B
CURRENT FILING DATE: 2005-02-14
PRIOR APPLICATION NUMBER: 60/544,190
PRIOR FILING DATE: 2004-02-13
NUMBER OF SEQ ID NOS: 119966
SEQ ID NO 82525
LENGTH: 334
TYPE: prt
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: peptide
LOCATION: (1)..(334)
OTHER INFORMATION: Ceres Seq. ID no. 12660048
US-11-056-355B-82525

Query Match 76.3%; Score 29; DB 7; Length 334;
Best Local Similarity 57.1%; Pred. No. 1.7e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
DB 306 WMLAAFT 312

```
RESULT 15
US-10-511-455-26
; Sequence 26, Application US/10511455
; Publication No. US20060088835A1
; GENERAL INFORMATION:
; APPLICANT: Pickard, Benjamin Simon
; APPLICANT: Blackwood, Douglas
; APPLICANT: Porteous, David
; APPLICANT: Muir, Walter John
; APPLICANT: Mors, Ole
; APPLICANT: Ewald, Henrik Lykke
; TITLE OF INVENTION: SCHIZOPHRENIA ASSOCIATED GENES
; FILE REFERENCE: 9013.63
; CURRENT APPLICATION NUMBER: US/10/511.455
; CURRENT FILING DATE: 2004-10-05
; PRIOR APPLICATION NUMBER: PCT/GB03/001543
; PRIOR FILING DATE: 2003-04-07
; PRIOR APPLICATION NUMBER: GB0207902.8
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: GB0207904.4
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: GB0207900.2
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: GB0207901.0
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: GB0227734.1
; PRIOR FILING DATE: 2002-11-28
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-511-455-26

Query Match      76.3%; Score 29; DB 6; Length 335;
Best Local Similarity 83.3%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      1 WMLSAF 6
      |||||
Db      316 WMLSI 321

RESULT 16
US-11-395-249-4
; Sequence 4, Application US/11395249
; Publication No. US20060177904A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavallee, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Werberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Agostino, Michael J.
; APPLICANT: Steindinger II, Robert J.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fechtel, Kim
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: 00766.000103.6
; CURRENT APPLICATION NUMBER: US/11/395,249
; CURRENT FILING DATE: 2006-04-03
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Homo sapiens
```

```
US-11-395-249-4

Query Match      76.3%; Score 29; DB 7; Length 335;
Best Local Similarity 83.3%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      1 WMLSAF 6
      |||||
Db      316 WMLSI 321

RESULT 17
US-11-056-355B-82529
; Sequence 82529, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nikolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 82529
; LENGTH: 338
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(338)
; OTHER INFORMATION: Ceres Seq. ID no. 1266052
US-11-056-355B-82529

Query Match      76.3%; Score 29; DB 7; Length 338;
Best Local Similarity 57.1%; Pred. No. 1.7e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY      1 WMLSAFS 7
      |||||
Db      310 WMLNAFT 316

RESULT 18
US-10-953-349-33299
; Sequence 33299, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 33299
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-10-953-349-33299

Query Match      76.3%; Score 29; DB 6; Length 345;
Best Local Similarity 71.4%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY      1 WMLSAFS 7
      |||||
Db      270 WMLIAFS 276
```

```
RESULT 19
US-10-953-349-38416
; Sequence 38416, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 38416
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-10-953-349-38416

Query Match          76.3%; Score 29; DB 6; Length 345;
Best Local Similarity 71.4%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 WMLSAFS 7
Db      270 WMLIAFS 276

RESULT 20
US-11-056-355B-79382
; Sequence 79382, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nickolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; TITLE OF INVENTION: Polypeptides Encoded Thereby
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 79382
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(351)
; OTHER INFORMATION: Ceres Seq. ID no. 12651073
US-11-056-355B-79382

Query Match          76.3%; Score 29; DB 7; Length 351;
Best Local Similarity 71.4%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 WMLSAFS 7
Db      273 WMLIAFS 279

RESULT 21
US-10-449-902-30511
; Sequence 30511, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
```

```
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30511
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-30511

Query Match          76.3%; Score 29; DB 6; Length 354;
Best Local Similarity 71.4%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 WMLSAFS 7
Db      279 WMLIAFS 285

RESULT 22
US-10-449-902-49957
; Sequence 49957, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 49957
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-49957

Query Match          76.3%; Score 29; DB 6; Length 354;
Best Local Similarity 71.4%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 WMLSAFS 7
Db      279 WMLIAFS 285

RESULT 23
US-10-953-349-32815
; Sequence 32815, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 32815
; LENGTH: 355
; TYPE: PRT
```

; ORGANISM: Zea mays subsp. mays
US-10-953-349-32815

Query Match
Best Local Similarity 76.3%; Score 29; DB 6; Length 355;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 WMLSAFS 7
Db 280 WMLTAFS 286

RESULT 24

US-11-056-355B-11231
; Sequence 11231, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:

APPLICANT: Brover, Vyacheslav
APPLICANT: Alexandrov, Nikolai
TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
FILE REFERENCE: 2750-1590PUS2
CURRENT APPLICATION NUMBER: US/11/056,355B
PRIOR FILING DATE: 2005-02-14
PRIOR APPLICATION NUMBER: 60/544,190
NUMBER OF SEQ ID NOS: 119966

SEQ ID NO 11231
LENGTH: 355
TYPE: PRT
ORGANISM: Zea mays subsp. mays
FEATURE:
NAME/KEY: peptide
LOCATION: (1)..(355)
OTHER INFORMATION: Ceres Seq. ID no. 13492044

US-11-056-355B-11231

Query Match
Best Local Similarity 76.3%; Score 29; DB 7; Length 355;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 WMLSAFS 7
Db 280 WMLTAFS 286

RESULT 25

US-11-056-355B-17746
; Sequence 17746, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:

APPLICANT: Brover, Vyacheslav
APPLICANT: Alexandrov, Nikolai
TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
FILE REFERENCE: 2750-1590PUS2
CURRENT APPLICATION NUMBER: US/11/056,355B
PRIOR FILING DATE: 2005-02-14
PRIOR APPLICATION NUMBER: 60/544,190
NUMBER OF SEQ ID NOS: 119966

SEQ ID NO 17746
LENGTH: 355
TYPE: PRT
ORGANISM: Zea mays subsp. mays
FEATURE:
NAME/KEY: peptide
LOCATION: (1)..(355)
OTHER INFORMATION: Ceres Seq. ID no. 13492044

US-11-056-355B-17746

Query Match
Best Local Similarity 76.3%; Score 29; DB 7; Length 355;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 WMLSAFS 7
Db 280 WMLTAFS 286

RESULT 26

US-10-953-349-33298
; Sequence 33298, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:

APPLICANT: ALEXANDROV, Nikolai et al.
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
FILE REFERENCE: 2750-1579PUS2
CURRENT APPLICATION NUMBER: US/10/953,349
CURRENT FILING DATE: 2004-09-30
NUMBER OF SEQ ID NOS: 40252
SOFTWARE: PatentIn version 3.3
SEQ ID NO 33298
LENGTH: 357
TYPE: PRT

US-10-953-349-33298

Query Match
Best Local Similarity 76.3%; Score 29; DB 6; Length 357;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 WMLSAFS 7
Db 282 WMLTAFS 288

RESULT 27

US-10-953-349-38415
; Sequence 38415, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:

APPLICANT: ALEXANDROV, Nikolai et al.
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
FILE REFERENCE: 2750-1579PUS2
CURRENT APPLICATION NUMBER: US/10/953,349
CURRENT FILING DATE: 2004-09-30
NUMBER OF SEQ ID NOS: 40252
SOFTWARE: PatentIn version 3.3
SEQ ID NO 38415
LENGTH: 357
TYPE: PRT

US-10-953-349-38415

Query Match
Best Local Similarity 76.3%; Score 29; DB 6; Length 357;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 WMLSAFS 7
Db 282 WMLTAFS 288

RESULT 28

US-11-330-403-6594
; Sequence 6594, Application US/11330403
; Publication No. US20060159563A1
; GENERAL INFORMATION:

APPLICANT: Abad, Mark S.
TITLE OF INVENTION: Genes and Uses for Plant Improvement
FILE REFERENCE: 38-21(53629)B
CURRENT APPLICATION NUMBER: US/11/330,403
CURRENT FILING DATE: 2006-01-12

NUMBER OF SEQ ID NOS: 19250
; SEQ ID NO 6594
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Magnaporthe grisea 70-15
US-11-330-403-6594

Query Match 76.3%; Score 29; DB 7; Length 367;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAF 6
|:|:|:|
Db 43 WMLSLF 48

RESULT 29
US-10-449-902-38671
; Sequence 38671, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A020571-US
; CURRENT APPLICATION NUMBER: US/10/449, 902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 38671
; LENGTH: 388
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-38671

Query Match 76.3%; Score 29; DB 6; Length 388;
Best Local Similarity 71.4%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
|:|:|:|
Db 311 WMLIAFS 317

RESULT 30
US-10-953-349-33297
; Sequence 33297, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 33297
; LENGTH: 417
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-10-953-349-33297

Query Match 76.3%; Score 29; DB 6; Length 417;
Best Local Similarity 71.4%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
|:|:|:|
Db 342 WMLIAFS 348

RESULT 31
US-10-953-349-32813
; Sequence 32813, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 32813
; LENGTH: 418
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-10-953-349-32813

Query Match 76.3%; Score 29; DB 6; Length 418;
Best Local Similarity 71.4%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
|:|:|:|
Db 343 WMLIAFS 349

RESULT 32
US-11-056-355B-11229
; Sequence 11229, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nickolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 11966
; SEQ ID NO 11229
; LENGTH: 418
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)-(418)
; OTHER INFORMATION: Ceres Seq. ID no. 13492042
US-11-056-355B-11229

Query Match 76.3%; Score 29; DB 7; Length 418;
Best Local Similarity 71.4%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
|:|:|:|
Db 343 WMLIAFS 349

RESULT 33
US-11-056-355B-17744
; Sequence 17744, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav

```
; APPLICANT: Alexandrov, Nickolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; PRIOR FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 17744
; LENGTH: 418
; TYPE: prt
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(418)
; OTHER INFORMATION: Ceres Seq. ID no. 13492042
US-11-056-355B-17744
```

```
Query Match      76.3%; Score 29; DB 7; Length 418;
Best Local Similarity 71.4%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 WMLSAFS 7
Db      343 WMLAFT 349
```

```
RESULT 34
US-11-056-355B-82524
; Sequence 82524, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nickolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 82524
; LENGTH: 437
; TYPE: prt
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(437)
; OTHER INFORMATION: Ceres Seq. ID no. 12666047
US-11-056-355B-82524
```

```
Query Match      76.3%; Score 29; DB 7; Length 437;
Best Local Similarity 57.1%; Pred. No. 2.2e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 WMLSAFS 7
Db      409 WMLAFT 415
```

```
RESULT 35
US-11-056-355B-82528
; Sequence 82528, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nickolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
```

```
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 82528
; LENGTH: 438
; TYPE: prt
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(438)
; OTHER INFORMATION: Ceres Seq. ID no. 12666051
US-11-056-355B-82528
```

```
Query Match      76.3%; Score 29; DB 7; Length 438;
Best Local Similarity 57.1%; Pred. No. 2.2e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 WMLSAFS 7
Db      410 WMLAFT 416
```

```
RESULT 36
US-10-953-349-31757
; Sequence 31757, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 31757
; LENGTH: 495
; TYPE: prt
; ORGANISM: Trifolium aestivum
US-10-953-349-31757
```

```
Query Match      76.3%; Score 29; DB 6; Length 495;
Best Local Similarity 71.4%; Pred. No. 2.3e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 WMLSAFS 7
Db      60 WMLAFT 66
```

```
RESULT 37
US-11-056-355B-66755
; Sequence 66755, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nickolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 66755
; LENGTH: 495
; TYPE: prt
; ORGANISM: Trifolium aestivum
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(495)
```


OTHER INFORMATION: Ceres Seq. ID no. 14302422
US-11-056-355B-66755

Query Match 76.3%; Score 29; DB 7; Length 495;
Best Local Similarity 71.4%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
|:|:|:|
DB 60 WALAAPS 66

RESULT 38
US-10-953-349-31756

; Sequence 31756, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 31756
; LENGTH: 530
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-953-349-31756

Query Match 76.3%; Score 29; DB 6; Length 530;
Best Local Similarity 71.4%; Pred. No. 2.7e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
|:|:|:|
DB 95 WALAAPS 101

RESULT 39
US-11-056-355B-66754
; Sequence 66754, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nickolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 66754
; LENGTH: 530
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(530)
; OTHER INFORMATION: Ceres Seq. ID no. 14302421
US-11-056-355B-66754

Query Match 76.3%; Score 29; DB 7; Length 530;
Best Local Similarity 71.4%; Pred. No. 2.7e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
|:|:|:|
DB 95 WALAAPS 101

RESULT 40
US-11-056-355B-82523

; Sequence 82523, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nickolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 82523
; LENGTH: 564
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(564)
; OTHER INFORMATION: Ceres Seq. ID no. 1266046
US-11-056-355B-82523

Query Match 76.3%; Score 29; DB 7; Length 564;
Best Local Similarity 57.1%; Pred. No. 2.9e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
|:|:|:|
DB 536 WMLAAPT 542

RESULT 41
US-11-056-355B-82527
; Sequence 82527, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nickolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 82527
; LENGTH: 565
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(565)
; OTHER INFORMATION: Ceres Seq. ID no. 1266050
US-11-056-355B-82527

Query Match 76.3%; Score 29; DB 7; Length 565;
Best Local Similarity 57.1%; Pred. No. 2.9e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 WMLSAFS 7
|:|:|:|
DB 537 WMLAAPT 543

RESULT 42
US-10-449-902-47030
; Sequence 47030, Application US/10449902
; Publication No. US20060123505A1

```
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A02051-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; PRIOR FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 47030
; LENGTH: 599
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-47030
```

```
Query Match          76.3%; Score 29; DB 6; Length 599;
Best Local Similarity 71.4%; Pred. No. 3e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 WMLSAFS 7
        |||
Db      168 WALAAPS 174
```

```
RESULT 43
US-10-953-349-31755
; Sequence 31755, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 31755
; LENGTH: 610
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-953-349-31755
```

```
Query Match          76.3%; Score 29; DB 6; Length 610;
Best Local Similarity 71.4%; Pred. No. 3.1e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 WMLSAFS 7
        |||
Db      175 WALAAPS 181
```

```
RESULT 44
US-11-056-355B-66753
; Sequence 66753, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nikolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
```

```
; SEQ ID NO 66753
; LENGTH: 610
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(610)
; OTHER INFORMATION: Ceres Seq. ID no. 14302420
US-11-056-355B-66753
```

```
Query Match          76.3%; Score 29; DB 7; Length 610;
Best Local Similarity 71.4%; Pred. No. 3.1e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 WMLSAFS 7
        |||
Db      175 WALAAPS 181
```

```
RESULT 45
US-10-953-349-33831
; Sequence 33831, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 33831
; LENGTH: 638
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-10-953-349-33831
```

```
Query Match          76.3%; Score 29; DB 6; Length 638;
Best Local Similarity 83.3%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 WMLSAF 6
        |||
Db      483 WMLSRF 488
```

```
RESULT 46
US-10-953-349-33830
; Sequence 33830, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 33830
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-10-953-349-33830
```

```
Query Match          76.3%; Score 29; DB 6; Length 665;
Best Local Similarity 83.3%; Pred. No. 3.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 WMLSAF 6
        |||
Db      510 WMLSRF 515
```

```
RESULT 47
US-10-953-349-33829
; Sequence 33829, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 33829
; LENGTH: 667
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-10-953-349-33829

Query Match      76.3%; Score 29; DB 6; Length 667;
Best Local Similarity 83.3%; Pred. No. 3.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 WMLSAF 6
Db      512 WMLSRF 517

RESULT 48
US-11-056-355B-1496
; Sequence 1496, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nikolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 1496
; LENGTH: 105
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(105)
; OTHER INFORMATION: Ceres Seq. ID no. 12411161
US-11-056-355B-1496

Query Match      73.7%; Score 28; DB 7; Length 105;
Best Local Similarity 66.7%; Pred. No. 90;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 WMLSAF 6
Db      100 WMLAAF 105

RESULT 49
US-11-056-355B-1495
; Sequence 1495, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nikolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 1495
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(124)
; OTHER INFORMATION: Ceres Seq. ID no. 12411160
US-11-056-355B-1495

Query Match      73.7%; Score 28; DB 7; Length 124;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 WMLSAF 6
Db      119 WMLAAF 124

RESULT 50
US-11-056-355B-1494
; Sequence 1494, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nikolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 1494
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(154)
; OTHER INFORMATION: Ceres Seq. ID no. 12411159
US-11-056-355B-1494

Query Match      73.7%; Score 28; DB 7; Length 154;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 WMLSAF 6
Db      149 WMLAAF 154

Search completed: August 29, 2006, 11:29:12
Job time : 5.54455 secs
```

THIS PAGE LEFT BLANK